

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:04:17 ; Search time 37.848 Seconds

(without alignments)  
4223.507 Million cell updates/sec

Title: US-10-014-269-34

Perfect score: 5279

Sequence: 1 MGEEGSGASHDEERASVTL.....LKINNCITYLGAELIAP 1007

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 156726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5279	100.0	1007	23	AAE25434 Human Nod2 mutant.
2	5279	100.0	1007	23	ABJ04766 Nod2a protein del
3	5272	99.9	1038	23	ABJ04735 Nod2a protein wild
4	5272	99.9	1040	23	AAE25416 Human Nod2 protein
5	5271	99.8	1007	23	ABJ04767 Nod2a P2685 + Iram
6	5271	99.8	1040	23	ABJ04773 Nod2a V9531 protei
7	5268	99.8	1040	23	ABJ04771 Nod2a V793K protei
8	5268	99.8	1040	23	ABJ04774 Nod2a N863V protei
9	5267	99.8	1040	23	AAE25418 Human Nod2 protein
10	5267	99.8	1040	23	ABJ04737 Nod2 protein sequ
11	5267	99.8	1040	23	ABJ04772 Nod2a N852S protei
12	5264	99.7	1040	23	ABJ04768 Nod2a R702W protei
13	5264	99.7	1040	23	ABJ04769 Nod2a G908R (SNP17
14	5264	99.7	1040	23	ABJ04770 Nod2a P2685 + N852
15	5259	99.6	1040	23	ABJ04776 Nod2a P2685 + N852
16	5256	99.6	1040	23	ABJ04775 Nod2a P2685 + R702
17	5256	99.6	1040	23	ABJ04777 Amino acid sequenc
18	5155	97.7	1041	22	AAE25417 Human Nod2 protein
19	5128	97.1	1013	23	ABJ04736 Nod2a protein wild
20	5128	97.1	1013	23	ABJ04736 Human CARD3X prote
21	4935.5	93.5	1009	23	AAU80865 Human CARD3X prote
22	4188	79.3	795	23	AAU80875 Human CARD3X prote
23	2959	56.1	560	23	AAU80856 Human CARD3X prote
24	2087	39.5	403	23	AAU80870 Human CARD3X prote
25	1614	30.6	305	23	AAE25421 Human Nod2 protein
26	1614	30.6	305	23	ABJ04740 Nod2 nucleotide bl
27	1564	29.6	296	23	AAU80869 Human CARD3X NB-AR
28	1175.5	22.3	953	20	AAJ31145 Murine CARD-4L (lon
29	1175.5	22.3	953	22	AAE20064 Mouse caspase recr
30	1175.5	22.3	953	23	ABG31080 Protein of murine
31	1175.5	22.3	953	23	AAO22136 Mouse Caspase recr
32	1175.5	22.3	953	24	ABU56298 Human protein SEQ
33	1173	22.2	953	21	AAE15552 Apoptosis related
34	1170.5	22.2	953	22	AAW78489 Human CARD-4L (lon
35	1170.5	22.2	953	22	AAE20080 Human Caspase recr
36	1170.5	22.2	953	22	AAO22111 Human Caspase recr
37	1170.5	22.2	953	24	ABU56273 Human CARD-4L (lon
38	1169.5	22.2	953	23	ABG31076 Human Caspase recr
39	1169.5	22.2	953	20	AAV31141 Human CARD-4L prot
40	1132	21.4	966	24	ABU56300 Caspase recruitmen
41	962	18.2	705	21	AAE15559 Apoptosis related
42	962	18.2	705	21	ABJ15562 Fragment of apopto
43	962	18.2	719	22	AAE75589 Human colon cancer
44	887.5	16.8	779	22	AAE95610 Human protein, sequ

Search completed: January 7, 2004, 20:49:55  
Job time : 41.848 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:54:33 ; Search time 29.6176 Seconds

(without alignments)  
6851.332 Million cell updates/sec

Title: US-10-014-269-34

Perfect score: 5279  
Sequence: 1 MGEEGGSASHDEERASVLT.....LKINSNCITYAGNELIQAAP 1007

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCONB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCONB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCONB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCONB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCONB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCONB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCONB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCONB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCONB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCONB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCONB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query  
No. Score Match Length DB ID Description

Search completed: January 7, 2004, 21:06:03  
Job time : 32.6176 secs

1	5279	100.0	1007	12	US-10-314-506-34	Sequence 34, Appl
2	5279	100.0	1007	14	US-10-014-269-34	Sequence 34, Appl
3	5279	100.0	1007	14	US-10-002-974-34	Sequence 34, Appl
4	5272	99.9	1040	12	US-10-314-506-2	Sequence 2, Appl
5	5272	99.9	1040	14	US-10-014-269-2	Sequence 2, Appl
6	5272	99.9	1040	14	US-10-002-974-2	Sequence 2, Appl
7	5271	99.8	1007	12	US-10-314-506-55	Sequence 55, Appl
8	5271	99.8	1007	14	US-10-002-974-55	Sequence 55, Appl
9	5271	99.8	1040	14	US-10-002-974-63	Sequence 63, Appl
10	5268	99.8	1040	14	US-10-002-974-63	Sequence 63, Appl
11	5268	99.8	1040	12	US-10-002-974-63	Sequence 63, Appl
12	5267	99.8	1040	14	US-10-314-506-4	Sequence 4, Appl
13	5267	99.8	1040	14	US-10-002-974-4	Sequence 4, Appl
14	5267	99.8	1040	14	US-10-002-974-4	Sequence 4, Appl
15	5267	99.8	1040	14	US-10-002-974-65	Sequence 65, Appl
16	5264	99.7	1040	12	US-10-314-506-57	Sequence 57, Appl
17	5264	99.7	1040	14	US-10-002-974-57	Sequence 57, Appl
18	5264	99.7	1040	14	US-10-002-974-59	Sequence 59, Appl
19	5264	99.7	1040	14	US-10-002-974-61	Sequence 61, Appl
20	5259	99.6	1040	14	US-10-002-974-87	Sequence 87, Appl
21	5256	99.6	1040	12	US-10-314-506-59	Sequence 59, Appl
22	5256	99.6	1040	12	US-10-314-506-61	Sequence 61, Appl
23	5256	99.6	1040	14	US-10-002-974-85	Sequence 85, Appl
24	5256	99.6	1040	14	US-10-002-974-89	Sequence 89, Appl
25	5155	97.7	1041	12	US-10-240-046A-2	Sequence 2, Appl
26	5128	97.1	1013	12	US-10-314-506-3	Sequence 3, Appl
27	5128	97.1	1013	14	US-10-014-269-3	Sequence 3, Appl
28	5128	97.1	1013	14	US-10-002-974-3	Sequence 3, Appl
29	4935.5	93.5	1009	10	US-09-864-921-107	Sequence 107, App
30	4188	79.3	795	10	US-09-864-921-188	Sequence 188, App
31	2959	56.1	560	10	US-09-864-921-83	Sequence 83, Appl
32	2087	39.5	403	10	US-09-864-921-176	Sequence 176, App
33	1614	30.6	305	12	US-10-314-506-7	Sequence 7, Appl
34	1614	30.6	305	14	US-10-014-269-7	Sequence 7, Appl
35	1614	30.6	305	14	US-10-002-974-7	Sequence 7, Appl
36	1564	29.6	296	10	US-09-864-921-174	Sequence 174, App
37	1175.5	22.3	953	9	US-09-728-721-43	Sequence 43, Appl
38	1175.5	22.3	953	14	US-10-118-984-43	Sequence 43, Appl
39	1175.5	22.3	953	15	US-10-295-981-43	Sequence 43, Appl
40	1170.5	22.2	953	9	US-09-728-721-8	Sequence 8, Appl
41	1170.5	22.2	953	12	US-10-028-374-4	Sequence 4, Appl
42	1170.5	22.2	953	12	US-10-183-770-4	Sequence 4, Appl
43	1170.5	22.2	953	14	US-10-105-931-8	Sequence 8, Appl
44	1170.5	22.2	953	14	US-10-118-984-8	Sequence 8, Appl
45	1170.5	22.2	953	15	US-10-013-477-12	Sequence 12, Appl

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:49:58 ; Search time 16.7833 Seconds  
(without alignments)  
5770.121 Million cell updates/sec

Title: US-10-014-269-34

Perfect score: 5279

Sequence: 1 MESEGGSSASHDEERASVLT.....LKLSNNCTTYLGAALILQAP 1007

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	404	7.7	1130	2	A4863		MHC class II trans
	2	390	7.4	1152	2	T1725		hypothetical prote
	3	350.5	6.6	1111	2	A59000		matern protein (imp
	4	232.5	4.4	456	2	A31857		ribonuclease inh
	5	227	4.3	461	2	A31858		ribonuclease-angio
	6	213	4.0	506	2	A45841		T-complex-associat
	7	210	4.0	456	2	S20597		ribonuclease inh
	8	210	4.0	535	2	T48102		RAN GTPase activat
	9	210	4.0	1075	2	T31668		hypothetical prote
	10	207.5	3.9	631	2	C89243		protein F28C1.3 (i
	11	207.5	3.9	631	2	T21471		hypothetical prote
	12	207	3.9	1004	2	T31665		hypothetical prote
	13	206	3.9	533	2	T52063		ran GTPase-activat

14	197	3.7	789	2	T52067	hypothetical prote
15	191.5	3.6	1010	2	T36383	probable large ATP
16	190.5	3.6	545	2	T52068	RAN GTPase-activat
17	185.5	3.5	1447	2	T42628	neutrophil apoptos
18	184.5	3.5	483	2	S27880	Nasopressin recept
19	173.5	3.3	862	2	T36380	probable large ATP
20	158	3.0	1232	2	A53478	neutrophil apoptos
21	156	3.0	332	2	T52069	GTPase-activating
22	154	2.9	1119	2	AB2239	hypothetical prote
23	150.5	2.9	312	2	B97746	hypothetical prote
24	150.5	2.9	1096	2	A96607	protein disease re
25	146	2.8	1411	2	S44370	C29A4.3 protein -
26	145	2.7	998	2	T23427	hypothetical prote
27	144.5	2.7	661	2	T36381	hypothetical prote
28	143.5	2.7	567	2	U55300	probable large ATP
29	140	2.7	1121	2	T02764	RAN GTPase activat
30	133	2.5	589	2	T52070	myosin-I binding p
31	130	2.5	618	2	T48193	RNAI protein homol
32	128	2.4	589	2	A36883	hypothetical prote
33	127.5	2.4	603	2	U66128	RNAI homolog fugu
34	126.5	2.4	1347	2	A12043	insulin-like growt
35	125.5	2.4	1256	2	AB2042	hypothetical prote
36	124	2.3	1151	2	T30936	hypothetical prote
37	122	2.3	183	2	F97803	reverse transcript
38	121.5	2.3	386	2	S37691	hypothetical prote
39	121	2.3	518	2	D96512	ran GTPase activat
40	121	2.3	800	2	AB1125	hypothetical prote
41	120.5	2.3	849	2	C97303	Interleukin A (limp
42	120	2.3	662	2	S42799	garp precursor - h
43	120	2.3	1091	2	A58532	glial cell membran
44	119.5	2.3	680	2	T19939	hypothetical prote
45	119.5	2.3	808	2	B97303	hypothetical prote

Search completed: January 7, 2004, 20:53:39  
Job time : 18.7833 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:05:32 ; Search time 11.1899 Seconds

(Without alignments)  
4232.403 Million cell updates/sec

Title: US-10-014-269-34

Perfect score: 5279

Sequence: 1 MGEEGGSASHDEBERASVLT.....LKLNNCTIYLGALLQAP 1007

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5272	99.9	1040	1	CARF_HUMAN
2	4070	77.1	1020	1	CARF_MOUSE
3	1178.5	22.3	953	1	CARD_MOUSE
4	1170.5	22.2	953	1	CARD_MOUSE
5	511	9.7	1062	1	PRAT_HUMAN
6	488.5	9.3	1033	1	CISI_MOUSE
7	486.5	9.2	1034	1	CISI_MOUSE
8	417.5	7.9	1155	1	CZTA_MOUSE
9	404	7.7	1130	1	CZTA_HUMAN
10	390	7.4	1473	1	NAL1_HUMAN
11	385.5	7.3	854	1	PRAS_RAT
12	374.5	7.1	980	1	PRAS_HUMAN
13	374	7.0	994	1	NAL1_HUMAN
14	372	7.0	1062	1	NAL1_HUMAN
15	361.5	6.8	892	1	PRAS_HUMAN
16	360.5	6.8	843	1	PRAS_MOUSE
17	350.5	6.6	1111	1	MATE_MOUSE

18	335	6.3	1200	1	MATE_HUMAN	P59047 homo sapien
19	292	5.5	1033	1	PRAS_HUMAN	P59045 homo sapien
20	232.5	4.4	456	1	RINI_PIG	P10775 sus scrofa
21	227	4.3	460	1	RINI_HUMAN	P13489 homo sapien
22	223	4.2	1403	1	BIRF_MOUSE	Q9j1b6 mus musculus
23	221.5	4.2	1024	1	CARC_HUMAN	Q9npp4 homo sapien
24	214.5	4.1	1402	1	BIRG_MOUSE	Q9j1b3 mus musculus
25	210	4.0	456	1	RINI_RAT	P29315 rattus norv
26	206.5	3.9	1403	1	BIRB_MOUSE	Q9gwks mus musculus
27	185.5	3.5	1447	1	BIRB_MOUSE	Q9gwks mus musculus
28	185	3.5	1403	1	BIRB_MOUSE	Q9gwks mus musculus
29	157.5	3.0	1403	1	BIRB_MOUSE	Q9gwks mus musculus
30	157.5	3.0	1403	1	BIRB_MOUSE	Q9gwks mus musculus
31	154.5	2.9	586	1	REP1_MOUSE	Q9j1b6 mus musculus
32	146	2.8	1411	1	YK63_MOUSE	P34342 xenopus lae
33	143.5	2.7	587	1	REP1_HUMAN	P46060 homo sapien
34	128	2.4	589	1	REP1_MOUSE	P46061 mus musculus
35	127.5	2.4	603	1	ALS_MOUSE	P70389 mus musculus
36	121.5	2.3	386	1	RNAI_MOUSE	P41391 schizosacch
37	121	2.3	800	1	RNAI_MOUSE	P41391 schizosacch
38	120	2.3	662	1	RNAI_MOUSE	P41391 schizosacch
39	117.5	2.2	1057	1	CARF_HUMAN	Q14392 homo sapien
40	114.5	2.2	4128	1	PRAD_MOUSE	Q9bxb6 homo sapien
41	113.5	2.2	2594	1	7LES_DROVY	P97313 mus musculus
42	113	2.1	359	1	TMO1_HUMAN	P20806 drosophila
43	112	2.1	603	1	ALS_MOUSE	P23289 homo sapien
44	112	2.1	820	1	TREB_MARTU	P35859 rattus norv
45	111	2.1	356	1	PGS2_COTTA	P54910 eglobacteri

ALIGNMENTS

RESULT 1

ID	CARF_HUMAN	STANDARD;	PRT; 1040 AA.
AC	Q9HC29; Q96RH5; Q96RH6; Q96RH8;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Caspase recruitment domain protein 15 (Nod2 protein) (Inflammatory		
DE	box-like disease protein 1).		
GN	CARD15 OR NOD2 OR IBD1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID:9606;		
PN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), MLAGENESIS OF LYS-305, AND		
RP	VARIANT ARG-908.		
RC	TISSUE=B-reast;		
RX	PubMed=11087742;		
RA	Ogura Y., Inohara N., Benito A., Chen F.F., Yamoka S., Nunez G.,		
RT	"Nod2, a Nod1/Apaef-1 family member that is restricted to monocytes and		
RT	activates NF-kappaB";		
RL	J Biol. Chem. 276:4612-4616 (2001).		
RN	[2]		



SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS CD THR-140, ARG-157;  
 RP CYS-235; ARG-248; SER-268; ASN-289; VAL-301; TRP-311;  
 RP VAL-348; ARG-352; CYS-373; SER-414; LEU-431; VAL-432; LYS-441;  
 RP VAL-612; THR-612; TRP-684; TRP-702; CYS-703; CYS-713; GLY-725;  
 RP ARG-908; ASP-918; ASP-924 AND ILE-955, VARIANTS ULCERATIVE COLITIS  
 RP THR-140 AND THR-885, AND VARIANT MET-189.  
 RC TISSUE=leukocyte;  
 RX MEDLINE=21279172; PubMed=1385576;  
 RA Hugot J.-P., Chamallard M., Zouali H., Lesage S., Cezard J.-P.,  
 RA Belachne J., Almer S., Tyk C., O'Morain C.A., Gassull M., Binder V.,  
 RA Finkel Y., Cortot A., Modigliani R., Laurent-Pug P.,  
 RA Gower-Rousseau C., Macry J., Colombel J.-F., Sabaotou M., Thomas G.,  
 RT "Association of NOD2 leucine-rich repeat variants with susceptibility  
 RT to Crohn's disease.";  
 RL Nature 411:599-603(2001).  
 RN [3].  
 RP VARIANTS BS GLN-334; TRP-334 AND PHE-469.  
 RX MEDLINE=21419644; PubMed=11528384;  
 RA McCell-Richard C., Lesage S., Rybojad M., Pileur A.M.,  
 RA Manouvrier-Hanu S., Harter R., Chamallard M., Zouali H., Thomas G.,  
 RA Hugot J.-P.,  
 RT "CARD15 mutations in Blau syndrome.";  
 RL Nat. Genet. 29:19-20(2001).  
 CC -1- FUNCTION: Induces NF-kappaB via RICK (CARDIAX, RIP2) and IKK-  
 gamma. Confers responsiveness to intracellular bacterial  
 1lipopolysaccharides (LPS).  
 CC -1- SUBUNIT: Binds to RICK by CARD-CARD interaction.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event:Alternative initiation;  
 CC Comment:2 isoforms, I/Nod2 (shown here) and Z/Nod2b are  
 CC produced by alternative initiation. Both isoforms can activate  
 CC NF-kappaB. Isoform 1 is more abundant;  
 CC -1- TISSUE SPECIFICITY: Monocytes-specific.  
 CC -1- DISEASE: Defects in CARD15 are the cause of Blau syndrome (BS)  
 CC [MIM:186580], a rare autosomal dominant disorder characterized by  
 CC early-onset granulomatous arthritis, uveitis and skin rash.  
 CC -1- DISEASE: Defects in CARD15 are a cause of susceptibility to  
 CC Crohn's disease (CD) [MIM:266600], a form of remitting  
 CC inflammatory bowel disease. CD may involve any part of the  
 CC gastrointestinal tract, but most frequently the terminal ileum and  
 CC colon. Bowel inflammation is transmural and discontinuous. Crohn's  
 CC disease is commonly classified as autoimmune disease.  
 CC -1- DISEASE: Defects in CARD15 are a cause of susceptibility to  
 CC ulcerative colitis [MIM:191390], a chronic inflammatory bowel  
 CC disease. In ulcerative colitis, the inflammation is continuous and  
 CC limited to rectal and colonic mucosal layers. Ulcerative colitis  
 CC is commonly classified as autoimmune disease.  
 CC -1- SIMILARITY: Contains 2 CARD domains.  
 CC -1- SIMILARITY: Contains 1 NACHT domain.  
 CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; AF178930; AAG33677.1; -  
 DR EMBL; AF385089; AAK70867.1; -  
 DR EMBL; AF385090; AAK70868.1; -  
 DR EMBL; AJ303140; CAC42117.1; -  
 DR Genem; HNC:5331; CARD15.  
 DR MIM; 605956; -  
 DR MIM; 186580; -  
 DR MIM; 266600; -  
 DR MIM; 191390; -  
 DR InterPro; IPR001315; CARD.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR007091; LRR\_NN1nh.  
 DR InterPro; IPR007111; NACHT\_NTPase.  
 DR Pfam; PF00560; LRR; 2.  
 DR SMART; SMO0114; CARD; 1.  
 DR PROSITE; PSS0209; CARD; 2.  
 DR PROSITE; PSS0637; NACHT; 1.  
 KW ATP-binding; Repeat; Leucine-rich repeat; Disease mutation;  
 KW Alternative initiation; Polymorphism.  
 FT CHAIN 1 1040  
 FT FT  
 FT CHAIN 28 1040  
 FT FT  
 FT INIT\_MET 28 28  
 FT DOMAIN 26 122  
 FT DOMAIN 126 218  
 FT DOMAIN 293 618  
 FT NP\_BIND 299 306  
 FT REPEAT 786 812  
 FT REPEAT 814 837  
 FT REPEAT 842 865  
 FT REPEAT 926 945  
 FT REPEAT 954 977  
 FT REPEAT 982 1005  
 FT VARIANT 140 140  
 FT VARIANT 157 157  
 FT VARIANT 189 189  
 FT VARIANT 235 235  
 FT VARIANT 248 248  
 FT VARIANT 268 268  
 FT VARIANT 289 289  
 FT VARIANT 291 291  
 FT VARIANT 294 294  
 FT VARIANT 301 301

ISOFORM 2.  
 CARD 1.  
 CARD 2.  
 NACHT.  
 ATP (POTENTIAL).  
 LRR 1.  
 LRR 2.  
 LRR 3.  
 LRR 4.  
 LRR 5.  
 LRR 6.  
 A -> T (in CD and ulcerative colitis).  
 W -> R (in CD).  
 /FTid=VAR\_012666.  
 T -> M.  
 /FTid=VAR\_012667.  
 R -> C (in CD).  
 /FTid=VAR\_012668.  
 L -> R (in CD).  
 /FTid=VAR\_012669.  
 P -> S.  
 /FTid=VAR\_012670.  
 N -> S.  
 /FTid=VAR\_012671.  
 D -> N (in CD).  
 /FTid=VAR\_012672.  
 T -> S (in CD).  
 /FTid=VAR\_012673.  
 A -> V (in CD).

FT	VARIANT	311	311	R -> W (in CD and ulcerative colitis)
FT	VARIANT	334	334	R -> W (in CD and /FTid=VAR_012674.
FT	VARIANT	334	334	R -> Q (in BS).
FT	VARIANT	334	334	R -> W (in BS).
FT	VARIANT	348	348	R -> V (in CD).
FT	VARIANT	352	352	R -> R (in CD).
FT	VARIANT	373	373	R -> C (in CD).
FT	VARIANT	414	414	R -> S (in CD).
FT	VARIANT	431	431	R -> L (in CD).
FT	VARIANT	432	432	R -> V (in CD).
FT	VARIANT	441	441	R -> K (in CD).
FT	VARIANT	469	469	R -> F (in BS).
FT	VARIANT	612	612	R -> T (in CD).
FT	VARIANT	612	612	R -> V (in CD).
FT	VARIANT	684	684	R -> W (in CD).
FT	VARIANT	702	702	R -> M (risk factor for CD).
FT	VARIANT	703	703	R -> C (in CD and ulcerative colitis)
FT	VARIANT	713	713	R -> C (in CD).
FT	VARIANT	725	725	R -> G (in CD).
FT	VARIANT	755	755	R -> V (in CD and ulcerative colitis)
FT	VARIANT	758	758	R -> V (in CD).
FT	VARIANT	778	778	R -> K (in CD).
FT	VARIANT	793	793	R -> M (in CD).
FT	VARIANT	843	843	R -> K (in CD).
FT	VARIANT	853	853	R -> S (in CD).
FT	VARIANT	863	863	R -> V (in CD).
FT	VARIANT	885	885	R -> T (in ulcerative colitis).
FT	VARIANT	908	908	R -> R (in CD).
FT	VARIANT	918	918	R -> D (risk factor for CD).

Db 66: ALIQKPPHILQITAFIAGLISREHWGLABEQISEKALLRQAKRMCARSIRKFFH 720  
 QY 721 SIIPAAGGAKSVHAMPGEIWLISLVEQCEERLARRAAGLNVGHUKLIFCSVGPTECA 780  
 Db 721 SIIPAAGGAKSVHAMPGEIWLISLVEQCEERLARRAAGLNVGHUKLIFCSVGPTECA 780  
 QY 781 ALAFVQLHRRPVALQIDNSVGDIGVEQLPCGVCCKLYIRDNNTSDRGICKILICGL 840  
 Db 781 ALAFVQLHRRPVALQIDNSVGDIGVEQLPCGVCCKLYIRDNNTSDRGICKILICGL 840  
 QY 841 HCEQLQMLAFNNKLTIDGCAHSMAKLLACRQNFALRLANNYITAGAQVLAEGIRGTS 900  
 Db 841 HCEQLQMLAFNNKLTIDGCAHSMAKLLACRQNFALRLANNYITAGAQVLAEGIRGTS 900  
 QY 901 LQFLGFWGNRVGDGAQALAEALGDHOSLRWTSLVGNNTGSVGAQALAMLARVMLEEL 960  
 Db 901 LQFLGFWGNRVGDGAQALAEALGDHOSLRWTSLVGNNTGSVGAQALAMLARVMLEEL 960  
 QY 961 CLEENHLQDEGVCSLAEGIKRNSSLKILKLSNCCITYLGAEPALLQA 1006  
 Db 961 CLEENHLQDEGVCSLAEGIKRNSSLKILKLSNCCITYLGAEPALLQA 1006

Search completed: January 7, 2004, 20:50:47  
 Job time : 16.1839 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:45:37 ; Search time 33.5667 Seconds (without alignments)

7741.578 Million cell updates/sec

Title: US-10-014-269-34

Perfect score: 5279 1 WGEQGASASHDEERASVLL.....LKLSSNCCITYLGAEPALLQA 1007

Sequence: 1 WGEQGASASHDEERASVLL.....LKLSSNCCITYLGAEPALLQA 1007

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP\_RP62L\_23:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_human:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mammal:\*
- 8: sp\_mhc:\*
- 9: sp\_organelle:\*
- 10: sp\_plant:\*
- 11: sp\_protent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
No.							

1	4070	77.1	1020	11	Q8K320	Q8K320 mus musculus
2	1178.5	22.3	953	11	Q8B3B0	Q8B3B0 mus musculus
3	1170.5	22.2	953	4	Q8IWF5	Q8IWF5 homo sapien
4	975	18.5	706	11	Q8BUL6	Q8BUL6 mus musculus
5	489.5	9.3	1033	11	Q8R4B8	Q8R4B8 mus musculus
6	479.5	9.1	892	4	Q8NEU4	Q8NEU4 homo sapien
7	419.5	7.9	1153	7	Q8CJEO	Q8CJEO rattus norv
8	417.5	7.9	1155	7	Q8TPE1	Q8TPE1 mus musculus
9	405.5	7.7	778	4	Q8NF48	Q8NF48 homo sapien
10	395.5	7.5	1052	7	Q8CJDB	Q8CJDB rattus norv
11	395.5	7.5	1073	7	Q8CJDB	Q8CJDB rattus norv
12	377.5	7.2	713	6	Q8S1Z7	Q8S1Z7 macaca fasc
13	357.5	6.8	733	4	Q8TEE2	Q8TEE2 homo sapien
14	354.5	6.7	982	11	Q8BU40	Q8BU40 mus musculus
15	339.5	6.4	1056	4	Q8NF42	Q8NF42 homo sapien
16	332	6.3	846	4	Q8IXT0	Q8IXT0 homo sapien
17	328	6.2	499	4	Q8TEI3	Q8TEI3 homo sapien
18	321	6.1	692	4	Q8GDS1	Q8GDS1 homo sapien
19	316.5	6.0	660	4	Q8NF06	Q8NF06 mus musculus
20	302	5.7	863	11	Q8C6J9	Q8C6J9 mus musculus
21	297	5.6	884	7	Q29675	Q29675 mus musculus
22	297	5.6	932	4	Q86K14	Q86K14 homo sapien
23	296.5	5.6	977	7	Q8HW99	Q8HW99 mus musculus
24	288	5.5	1097	4	Q8H6X0	Q8H6X0 mus musculus
25	282	5.3	312	11	Q8BV65	Q8BV65 mus musculus
26	282	5.3	397	11	Q8BVF4	Q8BVF4 mus musculus
27	268	5.1	605	10	Q93ZV8	Q93ZV8 arabidopsi
28	262	5.0	195	4	Q8NI01	Q8NI01 homo sapien
29	262	5.0	223	4	Q8NI02	Q8NI02 homo sapien
30	248	4.7	748	11	Q95WU0	Q95WU0 mus musculus
31	242.5	4.6	620	16	Q8XZN9	Q8XZN9 talstonia s
32	236	4.5	554	4	Q8HT24	Q8HT24 mus musculus
33	234.5	4.4	519	11	Q8C249	Q8C249 mus musculus
34	232.5	4.4	825	11	Q8CEM5	Q8CEM5 mus musculus
35	228	4.3	461	4	Q8IZK8	Q8IZK8 homo sapien
36	227	4.3	447	4	Q8FEF7	Q8FEF7 homo sapien
37	227	4.3	461	6	Q8HZB9	Q8HZB9 pan troglod
38	227	4.3	817	5	Q864V6	Q864V6 dictyostell
39	226.5	4.3	673	11	Q8CCN1	Q8CCN1 mus musculus
40	226	4.3	461	4	Q8BQ80	Q8BQ80 homo sapien
41	226	4.3	753	4	Q869L7	Q869L7 homo sapien
42	225	4.3	681	10	Q9XIK2	Q9XIK2 arabidopsi
43	223	4.2	1403	11	Q8CH64	Q8CH64 mus musculus
44	222	4.2	1403	11	Q8CH68	Q8CH68 mus musculus
45	216.5	4.1	1787	10	Q9M4X9	Q9M4X9 chlamydomon

Search completed: January 7, 2004, 20:52:41  
 Job time : 38.5667 secs

OM nucleic - protein search, using frame\_plus\_nbp model

Run on: January 7, 2004, 19:23:16 ; Search time 110.762 Seconds  
(without alignments)  
12857.211 Million cell updates/sec

Title: US-10-014-269-33  
Perfect score: 8270  
Sequence: 1 gtagcagatccagcagctaac.....ataacitgtgagtaaac 4486

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Command line parameters:

-MODEL=frame+\_nbp.model -DEV=xlp  
Q=/cgm2\_1/USPTO\_spool\_P/US10014269/runat\_07012004\_175228\_24497/app\_query.fasta\_1  
.9358  
-DB=A\_Geneseq\_19Jun03 -OFT=fastan -SUFFIX=tag -MINMATCH=0.1 -IOPTCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=150 -XCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50  
-MODE=LOCAL -OUTFM=pcr -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10014269 @CGN 1.1 311 @runat\_07012004\_175228\_24497 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONCLIC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

A\_Geneseq\_19Jun03:  
1: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1980.DAT:  
2: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1981.DAT:  
3: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1982.DAT:  
4: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1983.DAT:  
5: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1984.DAT:  
6: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1985.DAT:  
7: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1986.DAT:  
8: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1987.DAT:  
9: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1988.DAT:  
10: /SIDS1/gcdata/geneseq/geneseq-emb1/AA1989.DAT:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5435	65.7	1040	23	AAE25416 Human Nod2 protein
2	5434	65.7	1040	23	ABJ04773 Nod2a V935I protei
3	5432	65.7	1040	23	ABJ04771 Nod2a V935M protei
4	5431	65.7	1040	23	ABJ04774 Nod2a M663V protei
5	5430	65.7	1040	23	AAE25418 Human Nod2 protein
6	5430	65.7	1040	23	ABJ04737 Nod2 protein sequ
7	5430	65.7	1040	23	ABJ04772 Nod2a N855S protei
8	5427	65.6	1038	23	ABJ04735 Nod2a protein wild
9	5427	65.6	1040	23	ABJ04768 Nod2a R702W protei
10	5427	65.6	1040	23	ABJ04769 Nod2a P268S protei
11	5422	65.6	1040	23	ABJ04776 Nod2a P268S + N852
12	5419	65.5	1040	23	ABJ04775 Nod2a P268S + G980
13	5419	65.5	1040	23	ABJ04777 Nod2a P268S + R702
14	5419	65.5	1040	23	AAE25417 Amino acid sequen
15	5318	64.3	1041	22	AAE25417 Human Nod2 protein
16	5291	64.0	1013	23	ABJ04736 Human Nod2 protein
17	5279	63.8	1007	23	AAE25434 Human Nod2 mutant.
18	5279	63.8	1007	23	ABJ04766 Nod2a protein delt
19	5271	63.7	1007	23	ABJ04767 Human CARD3X prote
20	5271	63.7	1007	23	AAU80865 Human CARD3X prote
21	4981.5	60.2	1009	23	AAU80865 Human CARD3X prote
22	4188	50.6	795	23	AAU80875 Human CARD3X prote
23	2959	35.8	560	23	AAU80870 Human CARD3X prote
24	2087	25.2	403	23	AAE25421 Human Nod2 protein
25	1614	19.5	305	23	ABJ04740 Human Nod2 nucleotide bi
26	1614	19.5	305	23	AAU80869 Human CARD3X NB-AR
27	1564	18.9	296	23	AAV31145 Murine CARD-4L (lon
28	1206.5	14.6	953	22	AAE20080 Mouse caspase recr
29	1206.5	14.6	953	22	ABG31064 Protein of murine
30	1206.5	14.6	953	23	AAO22136 Mouse Caspase recr
31	1206.5	14.6	953	23	AAU80875 Human CARD3X prote
32	1206.5	14.6	953	23	AAU80875 Human CARD3X prote
33	1199	14.5	959	24	AAV79473 Human protein SEQ

34	1196.5	14.5	953	21	AAH15552	Apoptosis related	91	429	5.2	1130	17	AAH81569	Class II transacti
35	1196.5	14.5	953	22	AAH78489	Human protein S50	92	429	5.2	1130	19	AAW80313	Class II trans act
36	1196.5	14.5	953	23	AAH20080	Human CARD-4L (lon	93	429	5.2	1130	19	AAW57056	Class II trans act
37	1196.5	14.5	953	24	AAO22111	Protein of human C	94	425	5.1	320	23	ABJ04763	Nod1 protein sequ
38	1196.5	14.5	953	23	ABJ56273	Human Caspase recr	95	421.5	5.1	869	22	AAH61318	Human transactidlo
39	1195.5	14.5	953	23	ABG31076	Human caspase recr	96	418	5.1	81	23	AAH60668	Human CARD3X CARDB
40	1191.5	14.4	953	20	AAH31141	Human CARD-4L prot	97	416.5	5.0	896	23	AAO13592	Human CARD3X CARDB
41	1158	14.0	966	24	ABJ56300	Caspase recruitment	98	411	5.0	1397	22	AAV72670	Human PYRIN-11 pro
42	988	11.9	705	21	AAH15559	Apoptosis related	99	411	5.0	1442	22	AAV72671	Human NB-ARC and C
43	988	11.9	705	21	AAH15362	Fragment of apopto	100	411	5.0	1442	22	AAE06758	Human G-protein co
44	988	11.9	719	22	AAH75389	Human colon cancer	101	410.5	5.0	1429	23	ABG78455	Human caspase recr
45	887.5	10.7	779	22	AAH95610	Human protein sequ	102	410.5	5.0	1429	23	AAO17855	Human caspase recr
46	858	10.4	166	22	ABG49110	Human liver peptid	103	409	4.9	1424	22	AAV72674	Human caspase recr
47	858	10.4	166	22	ABG29107	Peptide #1758 enco	104	408	4.9	1424	22	AAV72674	Human caspase recr
48	858	10.4	166	22	ABJ19702	Protein #1701 enco	105	407	4.9	830	19	AAH80867	Human NAC gamma or
49	858	10.4	166	22	AAH55062	Human brain expres	106	407	4.9	1429	22	AAH62371	Human CARD3X CARDA
50	858	10.4	166	22	AAH67453	Human bone marrow	107	407	4.9	1429	22	ABG79769	Human CARD3X CARDA
51	858	10.4	166	22	AAH303024	Peptide #1706 enco	108	407	4.9	1429	23	ABG78472	Human CARD3X CARDA
52	858	10.4	166	22	ABG37074	Human peptide enco	109	407	4.9	1429	23	ABG79769	Human CARD3X CARDA
53	840	10.2	155	22	ABG38892	Human liver peptid	110	407	4.9	1429	23	ABG71631	Human CARD3X CARDA
54	840	10.2	155	22	ABG43475	Peptide #10981 enc	111	407	4.9	1429	24	ABG71631	Human CARD3X CARDA
55	840	10.2	155	22	ABG26439	Protein #6438 enco	112	407	4.9	1429	24	ABG71631	Human CARD3X CARDA
56	840	10.2	155	22	AAH64407	Human brain expres	113	407	4.9	1429	24	ABG71631	Human CARD3X CARDA
57	840	10.2	155	22	AAH77225	Human bone marrow	114	407	4.9	1429	24	ABG71631	Human CARD3X CARDA
58	840	10.2	155	22	AAH21157	Peptide #7591 enco	115	398.5	4.8	920	23	AAV72673	Human CARD3X CARDA
59	840	10.2	155	22	AAH7368	Peptide #11405 enc	116	396	4.8	920	23	AAV72673	Human CARD3X CARDA
60	840	10.2	155	22	ABG46236	Human peptide enco	117	396	4.8	920	23	AAV72673	Human CARD3X CARDA
61	820	9.9	146	22	ABG68853	Human liver peptid	118	395.5	4.8	919	23	ABP43483	Human CARD3X CARDA
62	820	9.9	146	22	ABH34746	Peptide #10982 enc	119	392	4.7	858	24	ABH13319	Human CARD3X CARDA
63	820	9.9	146	22	ABH6440	Protein #919 enco	120	392	4.7	858	24	ABH13319	Human CARD3X CARDA
64	820	9.9	146	22	AAH64408	Human brain expres	121	392	4.7	858	24	ABH13319	Human CARD3X CARDA
65	820	9.9	146	22	AAH77226	Human bone marrow	122	387	4.7	858	24	ABH13319	Human CARD3X CARDA
66	820	9.9	146	22	AAH21158	Peptide #7592 enco	123	387	4.7	858	24	ABH13319	Human CARD3X CARDA
67	820	9.9	146	22	AAH7369	Peptide #11406 enc	124	385.5	4.7	1022	23	AAO17862	Human CARD3X CARDA
68	820	9.9	146	22	ABG46237	Human peptide enco	125	384.5	4.6	952	22	ABG03924	Human CARD3X CARDA
69	639.5	7.7	180	23	AAH80876	Human CARD3X prote	126	382.5	4.6	1033	22	ABG07813	Human CARD3X CARDA
70	593	7.2	140	23	AAH80877	Mouse CARD3X #1.	127	382.5	4.6	1033	22	ABG07813	Human CARD3X CARDA
71	588	7.1	139	23	AAH80878	Mouse CARD3X #2.	128	382.5	4.6	1033	22	ABG07813	Human CARD3X CARDA
72	530.5	6.4	490	20	AAH31142	Human CARD-45 part	129	382	4.6	1062	22	AAH65895	Human CARD3X CARDA
73	530.5	6.4	490	22	AAH30081	Human CARD-45 (sho	130	382	4.6	1062	22	AAH65895	Human CARD3X CARDA
74	530.5	6.4	490	23	ABG31077	Human caspase recr	131	379.5	4.6	1049	22	AAH65895	Human CARD3X CARDA
75	530.5	6.4	490	23	AAO22127	Protein of human C	132	374.5	4.5	980	23	AAH65895	Human CARD3X CARDA
76	530.5	6.4	490	24	ABJ56289	Human Caspase recr	133	374.5	4.5	980	23	AAH65895	Human CARD3X CARDA
77	522	6.3	97	23	AAH25419	Human Nod2 protein	134	367	4.5	19938	24	AAH75525	Human CARD3X CARDA
78	522	6.3	97	23	ABJ04738	Nod2 caspase recr	135	366.5	4.4	891	22	AAH75525	Human CARD3X CARDA
79	521.5	6.3	1061	23	AAO15590	Human PYRIN-6 prot	136	366	4.4	740	22	ABG03831	Human CARD3X CARDA
80	521	6.3	1059	23	AAO17857	Pylin domain conta	137	360	4.4	19938	24	ABP76679	Human CARD3X CARDA
81	494.5	6.0	1034	22	AAE07514	Human PYRIN-1 prote	138	357	4.3	1111	23	AAU75525	Human CARD3X CARDA
82	494.5	6.0	1034	24	ABJ08503	Human PYRIN-1 prote	139	352	4.3	2312	23	AAO17858	Human CARD3X CARDA
83	485	5.9	94	23	AAE25420	Human Nod2 protein	140	351	4.2	1344	23	AAO15585	Human CARD3X CARDA
84	485	5.9	94	23	ABJ04739	Nod2 caspase recr	141	351	4.2	1344	23	AAO15585	Human CARD3X CARDA
85	439	5.3	975	22	ABP76796	Human cytokine alp	142	347.5	4.2	681	24	ABJ13320	Human CARD3X CARDA
86	438.5	5.3	19938	24	ABP76796	Streptomyces virid	143	346	4.2	19938	24	ABH98398	Human CARD3X CARDA
87	430	5.2	987	19	AAH76797	Class II trans act	144	344	4.2	1033	23	AAH65214	Human CARD3X CARDA
88	430	5.2	1207	19	AAH80311	Class II trans act	145	344	4.2	1033	23	AAH65214	Human CARD3X CARDA
89	429	5.2	1106	19	AAH80312	Class II transacti	146	344	4.2	1162	23	AAE2053	Human CARD3X CARDA
90	429	5.2	1130	16	AAH72452	Class II transacti	147	342.5	4.1	1200	23	AAH75526	Human CARD3X CARDA

148 342.5 4.1 1200 24 AAE31749  
149 342 4.1 1111 24 AAE31748  
150 341.3 4.1 777 14 AAE31740

Human WATER protei  
Mouse WATER protei  
Collagen-like poly

Search completed: January 7, 2004, 19:42:16  
Job time : 379.762 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model.

Run on: January 7, 2004, 19:49:02 / Search time 134.015 Seconds  
(without alignments)  
13490.627 Million cell updates/sec

Title: US-10-014-269-33  
Perfect score: 8270  
Sequence: 1 gtagacagatccaggtctcac.....ataaacgttgtagtaaac 4486

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 747907 seqs, 201509793 residues

Total number of hits satisfying chosen parameters: 1495814

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Command line parameters:

-MODEL=frame+\_n2p.model -DEV=xlp

Q=/cgn2\_1/USPTO\_spool\_p/US10014269/runat\_07012004\_175231\_24632/app\_query.fasta\_1  
.9358  
-DB=Published Applications AA -QFWT=faстан -SUFFIX=xapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=150 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=pct -NONW=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10014269\_@C@N\_1\_1\_59\_@runat\_07012004\_175231\_24632  
-NCRU=6 -ICPU=3 -NO\_WAP -LARGESQURY -NEG\_SCORES=0 -WAIT -DSPLOC=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

Published Applications AA: \*  
1: /cgn2\_6/ptodata/2/pubpa/US07\_PUBCOMB.pep: \*  
2: /cgn2\_6/ptodata/2/pubpa/PTI\_NEW\_PUB.pep: \*  
3: /cgn2\_6/ptodata/2/pubpa/US06\_NEW\_PUB.pep: \*  
4: /cgn2\_6/ptodata/2/pubpa/US06\_PUBCOMB.pep: \*  
5: /cgn2\_6/ptodata/2/pubpa/US07\_NEW\_PUB.pep: \*  
6: /cgn2\_6/ptodata/2/pubpa/PTUS\_PUBCOMB.pep: \*  
7: /cgn2\_6/ptodata/2/pubpa/US08\_NEW\_PUB.pep: \*  
8: /cgn2\_6/ptodata/2/pubpa/US08\_PUBCOMB.pep: \*  
9: /cgn2\_6/ptodata/2/pubpa/US09A\_PUBCOMB.pep: \*  
10: /cgn2\_6/ptodata/2/pubpa/US09B\_PUBCOMB.pep: \*





97	407	4.9	1429	12	US-10-183-770-15
98	407	4.9	1429	15	US-10-028-392-11
99	407	4.9	1454	10	US-09-388-221-10
100	407	4.9	1473	10	US-09-388-221-2
101	393.5	4.8	919	12	US-10-094-749-2718
102	392	4.7	858	9	US-09-848-035-8
103	392	4.7	858	10	US-09-866-224-8
104	392	4.7	994	11	US-09-965-621-16
105	392	4.7	994	12	US-10-407-866-16
106	392	4.7	994	15	US-10-066-521-24
107	384.5	4.6	952	12	US-10-407-866-70
108	382.5	4.6	1033	12	US-10-132-967-2
109	382.5	4.6	1033	14	US-10-127-516-2
110	382.5	4.6	1033	14	US-10-027-629-2
111	382	4.6	1062	12	US-10-239-663-43
112	379.5	4.6	1049	12	US-10-239-663-42
113	379	4.6	635	12	US-10-407-866-90
114	378.5	4.6	772	12	US-10-094-749-2366
115	374.5	4.5	980	9	US-09-848-035-13
116	374.5	4.5	980	10	US-09-965-621-13
117	363.5	4.4	947	11	US-09-965-621-18
118	363.5	4.4	947	12	US-10-407-866-18
119	362	4.4	902	12	US-10-407-866-66
120	357	4.3	1111	15	US-10-216-645-5
121	352.5	4.3	674	12	US-10-407-866-89
122	352.5	4.3	719	11	US-09-965-621-59
123	352.5	4.3	719	12	US-10-407-866-59
124	351	4.2	1344	15	US-10-066-521-6
125	350	4.2	612	12	US-10-407-866-88
126	344	4.2	1143	15	US-10-216-645-4
127	344	4.2	1162	15	US-10-342-331-4
128	341.5	4.1	720	12	US-10-342-331-4
129	339.5	4.1	635	12	US-10-296-539-2
130	339.5	4.1	635	12	US-10-407-866-84
131	339.5	4.1	635	15	US-10-066-521-13
132	339	4.1	590	12	US-10-407-866-94
133	338	4.1	1466	12	US-10-301-822-33
134	338	4.1	1466	15	US-10-177-293-68
135	336.5	4.1	720	12	US-10-342-331-4
136	331	4.0	584	12	US-10-029-866-32446
137	329.5	4.0	821	9	US-09-848-035-2
138	329.5	4.0	821	10	US-09-966-224-2
139	329	4.0	200	9	US-09-728-721-11
140	329	4.0	200	14	US-10-105-931-11
141	329	4.0	200	15	US-10-118-984-11
142	329	4.0	200	15	US-10-285-981-11
143	329.5	4.0	1078	15	US-10-058-124-21
144	324.5	3.9	960	12	US-10-342-331-5
145	321.5	3.9	321	12	US-10-407-866-116
146	321.5	3.9	659	12	US-10-407-866-91
147	316	3.8	875	9	US-09-848-035-16
148	316	3.8	875	10	US-09-966-224-16
149	316	3.8	4123	15	US-10-213-509-5
150	315.5	3.8	1456	12	US-10-301-822-35

Search completed: January 7, 2004, 20:54:10  
Job time : 430.015 secs

Sequence 15, Appl  
Sequence 11, Appl  
Sequence 10, Appl  
Sequence 2, Appl  
Sequence 2718, Ap  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 24, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 43, Appl  
Sequence 40, Appl  
Sequence 92, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 18, Appl  
Sequence 66, Appl  
Sequence 59, Appl  
Sequence 59, Appl  
Sequence 2, Appl  
Sequence 88, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 84, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 68, Appl  
Sequence 33, Appl  
Sequence 32446, A  
Sequence 2, Appl  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 21, Appl  
Sequence 116, Appl  
Sequence 91, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 35, Appl

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK nucleic - protein search, using frame\_plus\_n2p model

Run on: January 7, 2004, 19:24:31 / Search time 65.5073 seconds

(without alignments)  
13171.439 Million cell updates/sec

Title: US-10-014-269-33

Perfect score: 8270  
Sequence: 1 gtagacagatccaggtccac.....ataactgttgatcaaac 4486

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 56616

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Command line parameters:

-MODEL=frame+ntp.model -DEV=xlp  
Q=/cgm2\_1/USPTO\_spool\_p/US10014269/runat\_07012004\_175229\_24535/app\_query.fasta\_1  
-DB=PIR 76 -QPM=fastan -SUFFIX=prf -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40 cdi -LIST=150  
-DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTEXT=ptio -NOM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USFEUS10014269 -GCEN 1.1 224 -stimac 07012004\_175229\_24535 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -THREADS=1 -XGAPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 76:  
1: p1r1:  
2: p1r2:  
3: p1r3:  
4: p1r4:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	429	5.2	1130	2 A4B843	MHC class II trans
2	390	4.7	1192	2 T17255	hypothetical prote
3	357	4.3	1111	2 A59000	matern protein [imp
4	338	4.1	1466	1 CGHUL7	collagen alpha 1(I
5	329.5	4.0	1453	2 S21626	collagen alpha 1(I
6	322.5	3.9	2944	2 A54849	collagen alpha 1(I
7	313.5	3.8	1496	1 CGHUV2V	collagen alpha 1(I
8	312.5	3.8	1464	2 S59856	collagen alpha 1(I
9	312	3.8	784	2 J00317	hypothetical 82K p
10	310.5	3.8	1464	1 CGHUL7	collagen alpha 1(I
11	310.5	3.8	1466	1 S59856	collagen alpha 1(I
12	310	3.8	1464	2 S59856	collagen alpha 1(I
13	309	3.8	1373	1 A43291	collagen alpha 2(I
14	308.5	3.8	660	1 Q08E3	BHLF1 protein - hu
15	305	3.7	1049	1 CGH07S	collagen alpha 1(I
16	301.5	3.7	1487	1 CGH06C	collagen alpha 5(I
17	301.5	3.6	1691	1 S22917	collagen alpha 3(I
18	297.5	3.6	1670	1 CGH03B	collagen alpha 3(I
19	294.5	3.6	1049	1 CGH07S	collagen alpha 3(I
20	294.5	3.6	1418	2 T45467	collagen alpha 1(I
21	294	3.6	779	1 S23809	collagen alpha 1(I
22	291	3.5	1414	1 S23809	collagen alpha 1(I
23	289.5	3.5	2715	2 T13049	collagen alpha 2(I
24	288.5	3.5	1497	2 T13049	eyelid - fruit fly
25	288.5	3.5	1546	1 CGH02E	collagen alpha 2(I
26	287.5	3.5	1366	1 CGH02E	collagen alpha 2(I
27	287	3.5	886	2 T50694	collagen alpha 1(I
28	287	3.5	886	2 T50694	collagen alpha 1(I
29	287	3.5	886	2 T50694	collagen alpha 1(I
30	286.5	3.5	1690	1 CGH01B	collagen alpha 4(I
31	284	3.5	671	1 CGH01S	collagen alpha 1(I
32	283.5	3.5	1487	2 B41182	collagen alpha 1(I
33	283.5	3.5	1487	2 B41182	collagen alpha 1(I
34	282.5	3.4	2944	2 A54849	collagen alpha 1(I
35	281	3.4	1806	1 CGH01E	collagen alpha 1(I
36	280	3.4	1838	1 CGH01V	collagen alpha 1(I
37	276.5	3.4	1419	2 A41182	collagen alpha 1(I
38	275.5	3.4	1042	1 CGH01S	collagen alpha 1(I
39	273.5	3.3	1453	2 S21626	collagen alpha 1(I
40	269.5	3.3	1418	2 T45467	collagen alpha 1(I
41	269	3.3	1464	1 CGH01S	collagen alpha 1(I
42	268.5	3.3	1603	2 S23809	collagen alpha 1(I
43	268	3.2	1042	1 CGH01S	collagen alpha 1(I
44	268	3.2	1419	2 A41182	collagen alpha 1(I
45	268	3.2	1487	2 B41182	collagen alpha 1(I
46	268	3.2	1843	2 S18803	collagen alpha 1(I
47	266	3.2	1691	1 S22917	collagen alpha 1(I
48	264	3.2	1585	2 T31611	hypothetical prote
49	264	3.2	1669	1 CGH04B	collagen alpha 1(I
50	263.5	3.2	1496	1 CGH02V	collagen alpha 2(I
51	263.5	3.2	1606	1 CGH01E	collagen alpha 1(I
52	263	3.2	1669	1 CGH04B	collagen alpha 1(I
53	262.5	3.2	1492	2 A40333	collagen alpha 1(I
54	261.5	3.2	1487	1 CGH06C	collagen alpha 1(I
55	259	3.1	779	1 CGH01S	collagen alpha 1(I
56	258.5	3.1	1763	2 S16366	collagen alpha 2(I
57	258	3.1	436	2 A31697	ribonuclease inhib
58	257.5	3.1	2715	2 T13049	eyelid - fruit fly
59	257	3.1	1373	1 A43291	collagen alpha 2(I
60	256.5	3.1	1838	2 S28791	collagen alpha 1(I
61	255.5	3.1	255.5	1 CGH01V	collagen alpha 1(I
62	255.5	3.1	1492	2 A40333	collagen alpha 1(I
63	255.5	3.1	1546	1 CGH02E	collagen alpha 1(I
64	254.5	3.1	1670	1 CGH03B	collagen alpha 2(I
65	254.5	3.1	1843	2 S18803	collagen alpha 1(I
66	253	3.1	1366	1 CGH02S	collagen alpha 1(I
67	252	3.0	660	1 Q08E3	BHLF1 protein - hu
68	252	3.1	1486	1 B40333	collagen alpha 1(I
69	250	3.0	1486	1 B40333	collagen alpha 1(I
70	249.5	3.0	1758	2 T29350	hypothetical prote
71	249.5	3.0	1759	2 T29351	collagen alpha 2(I
72	247	3.0	1344	1 A35175	collagen alpha 2(I
73	246.5	3.0	920	2 B34493	collagen alpha 1(I
74	245.5	3.0	1603	2 S23810	collagen alpha 1(I
75	243.5	2.9	1357	2 T29265	hypothetical prote
76	243.5	2.9	1497	2 T29265	hypothetical prote
77	243.5	3.0	1549	2 T29265	hypothetical prote
78	241.5	2.9	1712	1 CGH02B	type VII collagen
79	241.5	2.9	1763	2 S16366	collagen alpha 2(I
80	241	2.9	744	1 A34246	collagen alpha 1(I
81	241	2.9	744	1 S23298	collagen alpha 1(I
82	240.5	2.9	964	1 CGH02S	collagen alpha 1(I
83	240	2.9	13288	2 T03099	collagen alpha 2(I
84	239.5	2.9	920	2 A45748	collagen alpha 1(I
85	239.5	2.9	1549	2 T29265	collagen alpha 1(I
86	239	2.9	1460	1 EDBE1F	type VII collagen
87	239	2.9	1707	2 A33526	immediate-early pr
88	238.5	2.9	627	2 A44112	collagen alpha 2(I
89	238	2.9	1758	2 T29350	hypothetical prote
90	238	2.9	1759	2 T29351	collagen alpha 2(I
91	238	2.9	1791	2 T02345	collagen alpha 2(I
92	237	2.9	1428	2 T08852	hypothetical prote
93	236.5	2.9	1106	2 J00405	lustrin A - Califo
94	235	2.8	931	2 S13380	hypothetical 119.5
95	234.5	2.8	4957	2 T03455	collagen alpha 1(I
96	234.5	2.8	1691	2 CGH05B	collagen alpha 1(I
97	234.5	2.8	5262	2 T03454	collagen alpha 6(I
98	234	2.8	1669	1 CGH05B	collagen alpha 6(I
99	233.5	2.8	1106	2 J00405	collagen alpha 1(I
100	232.5	2.8	506	2 A45841	hypothetical 119.5
101	231.5	2.8	754	2 A35267	T-complex-associat
102	231	2.8	920	2 A45748	collagen alpha 5(I
103	229.5	2.8	1585	2 T31611	collagen alpha 1(I
104	229.5	2.8	3198	2 A43426	hypothetical prote
105	229.5	2.8	825	2 J04163	collagen alpha 2 f
106	228	2.8	1446	1 A45344	DNA-binding protei
107	228	2.8	461	2 A31858	immediate-early pr
108	227.5	2.7	1360	2 T33922	ribonuclease-angio
109	226	2.7	1360	2 T33922	hypothetical prote
110	223.5	2.7	2142	2 B35098	MHC class III hist

c 111	225	2.7	1024	2	S18251	collagen alpha 1(X
c 112	225	2.7	1669	1	CGH48	collagen alpha 1(I
c 113	224	2.7	1184	2	S50832	atrophin-1 - huma
c 114	223.5	2.7	1027	2	S28774	collagen alpha cha
c 115	223.5	2.7	1574	2	T13954	MEGF6 protein - ra
c 116	223	2.7	1024	2	S18251	collagen alpha 1(X
c 117	222.5	2.7	1446	1	A45344	immediate-early pr
c 118	222	2.7	671	1	CGRT15	collagen alpha 1(I
c 119	222	2.7	1774	2	B6101	collagen alpha 1(X
c 120	222	2.7	4957	2	T03455	collagen alpha 1(X
c 121	222	2.7	5262	2	T03454	collagen alpha 2(I
c 122	221.5	2.7	964	1	CGCH25	collagen alpha 2(I
c 123	221.5	2.7	2321	2	S78549	notch3 protein - h
c 124	220.5	2.7	1315	2	A56101	collagen alpha 1(X
c 125	219.5	2.7	1712	1	CGH28	collagen alpha 2(I
c 126	219	2.6	456	2	S20597	ribonuclease inh
c 127	219	2.7	1752	2	A45407	collagen alpha 3(I
c 128	218.5	2.6	2232	2	T34434	hypothetical prote
c 129	218	2.6	677	2	S23296	collagen alpha 2(I
c 130	218	2.6	744	2	S15435	collagen alpha 1(X
c 131	218	2.7	1870	2	S37671	MHC class III hist
c 132	217.5	2.6	1184	2	G01763	atrophin-1 - huma
c 133	217.5	2.7	1872	2	S36152	MHC class III hist
c 134	216	2.6	13288	2	T03099	mucin, submaxillar
c 135	215.5	2.6	743	1	S23779	collagen alpha 1(I
c 136	215.5	2.6	921	2	S42617	collagen alpha 1(I
c 137	215.5	2.6	1532	2	A61262	collagen alpha 1(X
c 138	215	2.6	635	2	A57131	collagen alpha 2(I
c 139	214.5	2.6	635	2	A57131	collagen alpha 2(I
c 140	214.5	2.6	1574	2	T13954	MEGF6 protein - ra
c 141	214	2.6	1691	1	CGH48	collagen alpha 6(I
c 142	213.5	2.6	920	2	S34493	collagen alpha 1(I
c 143	213	2.6	931	2	S13580	collagen alpha 1(I
c 144	213	2.6	1460	1	EDBE1F	immediate-early pr
c 145	212.5	2.6	2796	2	JC4743	fatty-acid synthas
c 146	212	2.6	825	1	EDBE1D	immediate-early pr
c 147	211.5	2.6	1964	2	T09059	notch4 - mouse
c 148	211	2.6	730	2	A36226	collagen alpha 1 c
c 149	210.5	2.5	1027	2	S28774	collagen alpha cha
c 150	210.5	2.5	1744	2	S40991	collagen alpha 1(I

Search completed: January 7, 2004, 19:54:38  
Job time : 511.507 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 7, 2004, 19:24:01 ; Search time 34.7539 Seconds

(without alignments)  
12140.333 Million cell updates/sec

Title: US-10-014-269-33

Perfect score: 8270  
Sequence: 1 gttagacagatccagctcac.....ataaactgttagctcaaac 4486

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=framet\_n2p.model -DEV=xlp

Q=/cgn2\_1/USPTO\_spool\_p/US10014269/runac\_07012004\_175229\_24504/app\_query.fasta\_1

9356

-DB=SwissProt\_41 -QMT=fastan -SUFFIX=frsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humand0 cdi -LIST=150

-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL

-OUTFMT=psio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=us10014269\_gcn\_1.1.97 @runac\_07012004\_175229\_24504 -NCPU=6 -ICPU=3

-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	5435	65.7	1040	1	CARF_HUMAN	Q9h29	homo sapien	58	268	3.2	1453	1	CA11_CHICK	P02457	gallus gall
2	4183	50.6	1020	1	CARF_MOUSE	Q9K30	mus musculus	58	268	3.2	1459	1	CA12_MOUSE	P28481	mus musculus
3	1209.5	14.6	953	1	CAR4_MOUSE	Q9bhd0	mus musculus	59	267	3.3	1355	1	CA21_RANCA	O42350	rana catesb
4	1196.5	14.5	953	1	CAR4_HUMAN	Q9y229	homo sapien	60	267	3.3	1355	1	CA14_HUMAN	P02462	homo sapien
5	523	6.3	1062	1	PYA7_HUMAN	P59046	homo sapien	61	264	3.2	1669	1	CA14_MOUSE	P05397	homo sapien
6	495.5	6.0	1033	1	CIS1_MOUSE	Q9t488	mus musculus	62	263.5	3.2	1496	1	CA14_MOUSE	P02463	mus musculus
7	494.5	6.0	1034	1	CIS1_HUMAN	Q9620	homo sapien	63	263	3.2	1669	1	CA14_MOUSE	P29400	homo sapien
8	429	5.2	1130	1	C27A_HUMAN	P33076	homo sapien	64	262	3.2	1685	1	CA54_HUMAN	P29400	homo sapien
9	417.5	5.0	1155	1	C27A_MOUSE	P33076	homo sapien	65	261.5	3.2	1685	1	CA12_HUMAN	P02458	homo sapien
10	407	4.9	1473	1	NAL1_MOUSE	P79962	mus musculus	66	259.5	3.1	1362	1	CA21_CHICK	P02457	gallus gall
11	382	4.7	994	1	NAL1_HUMAN	Q9c000	homo sapien	67	259	3.1	1779	1	CA11_BOVIN	P02457	gallus gall
12	385.5	4.7	854	1	PYA5_RAT	Q9c305	rattus norv	68	258.5	3.1	1763	1	CA24_ASCU	P27393	ascaris suu
13	382	4.6	1062	1	NAL2_HUMAN	Q9h94	homo sapien	69	258	3.1	1456	1	RINI_PIG	P10775	sus scrofa
14	374.5	4.5	980	1	PYA3_HUMAN	Q9h94	homo sapien	70	258	3.1	1366	1	CA21_HUMAN	P08123	homo sapien
15	366	4.4	892	1	PYA5_HUMAN	P59044	homo sapien	71	257	3.1	1372	1	CA21_MOUSE	Q01149	mus musculus
16	360.5	4.4	843	1	PYA5_MOUSE	Q9t482	mus musculus	72	256.5	3.1	1806	1	CA1B_HUMAN	P12107	homo sapien
17	357	4.3	1111	1	MATE_MOUSE	Q9t482	mus musculus	73	255.5	3.1	1838	1	CA15_HUMAN	P20908	homo sapien
18	342.5	4.1	1200	1	MATE_HUMAN	P59047	homo sapien	74	255.5	3.1	1364	1	CA21_BOVIN	P02465	bos taurus
19	338	4.1	1466	1	MATE_HUMAN	P02461	homo sapien	75	254	3.1	1366	1	CA28_HUMAN	O46392	gallus gall
20	329.5	4.0	1453	1	CA13_HUMAN	P11087	mus musculus	76	252	3.0	1736	1	CA28_HUMAN	O46392	gallus gall
21	322.5	3.9	2944	1	CA17_HUMAN	Q02388	homo sapien	77	252	3.0	1660	1	YH11_BEV	P03181	epstein-bar
22	318	3.9	1464	1	CA13_MOUSE	P05397	homo sapien	78	252	3.0	1366	1	CA1B_MOUSE	P02459	bos taurus
23	315.5	3.8	1464	1	CA25_HUMAN	P02452	homo sapien	79	251.5	3.0	1804	1	CA1B_MOUSE	P02459	bos taurus
24	312.5	3.8	1464	1	CA11_HUMAN	P02452	homo sapien	80	250	3.0	1670	1	CA12_BOVIN	O46392	mus musculus
25	312	3.8	784	1	YAN2_HUMAN	P14728	xanthomomas	81	249.5	3.0	1747	1	CA12_BOVIN	O46392	mus musculus
26	310.5	3.8	1466	1	YAN2_XANCV	P02461	homo sapien	82	249.5	3.0	1758	1	CA24_CAEEL	P17140	caenorhabdi
27	309	3.7	1033	1	PYA6_HUMAN	P02461	homo sapien	83	249.5	3.0	1804	1	CA24_CAEEL	O01955	bos taurus
28	309	3.7	1372	1	CA21_MOUSE	Q01149	mus musculus	84	246.5	3.0	1603	1	CA1F_HUMAN	O01955	bos taurus
29	308.5	3.8	660	1	YH11_BEV	P03181	epstein-bar	85	244.5	3.0	1362	1	CA1F_HUMAN	O01955	bos taurus
30	307.5	3.7	1464	1	CA13_MOUSE	P08121	mus musculus	86	243.5	3.0	1747	1	CA12_BOVIN	O01955	bos taurus
31	305	3.7	1048	1	CA13_BOVIN	P04258	bos taurus	87	243.5	3.0	1285	1	MUC1_HUMAN	P15941	h. mucin 1 p
32	304.5	3.7	1460	1	CA11_CANFA	Q9x517	canis fami	88	242.5	2.9	1516	1	CA1H_HUMAN	P03060	homo sapien
33	301.5	3.7	1418	1	CA12_HUMAN	Q9x517	canis fami	89	242.5	2.9	1712	1	CA24_HUMAN	P03060	homo sapien
34	297.5	3.6	1685	1	CA54_HUMAN	P29400	homo sapien	90	241.5	2.9	1763	1	CA24_HUMAN	P03060	homo sapien
35	294.5	3.6	1049	1	CA13_BOVIN	P02458	bos taurus	91	241.5	2.9	1763	1	CA24_HUMAN	P03060	homo sapien
36	294	3.6	779	1	CA11_BOVIN	P02458	bos taurus	92	241.5	2.9	1763	1	CA24_HUMAN	P03060	homo sapien
37	292.5	3.5	1670	1	CA34_HUMAN	O01955	bos taurus	93	240.5	2.9	1763	1	CA24_HUMAN	P03060	homo sapien
38	290.5	3.5	1262	1	CA13_CHICK	P12105	gallus gall	94	239	2.9	1355	1	CA21_RANCA	P14282	oryctolagus
39	289	3.5	1366	1	CA21_HUMAN	P08123	homo sapien	95	239	2.9	2003	1	NTC4_HUMAN	O42350	rana catesb
40	288	3.5	1262	1	CA13_CHICK	P12105	gallus gall	96	238.5	2.9	1763	1	CA24_MOUSE	P08122	mus musculus
41	288	3.5	1736	1	CA2B_HUMAN	P13942	homo sapien	97	238	2.9	1758	1	SPD2_NEPCL	O99466	homo sapien
42	287	3.5	1690	1	CA44_HUMAN	P53420	homo sapien	98	237	2.9	1736	1	CA24_CAEEL	P17140	caenorhabdi
43	286.5	3.5	1690	1	CA44_HUMAN	P53420	homo sapien	99	236.5	2.9	1736	1	CA24_CAEEL	P17140	caenorhabdi
44	283.5	3.4	2944	1	CA17_HUMAN	P02465	bos taurus	100	236.5	2.9	1024	1	CA19_HUMAN	Q9h94	homo sapien
45	281.5	3.4	1364	1	CA21_BOVIN	P02465	bos taurus	101	235.5	2.8	1356	1	CA21_ONCVY	O93484	oncorhynch
46	281	3.4	671	1	CA11_RAT	P02453	rattus norv	102	235	2.8	1707	1	CA24_MOUSE	P08122	mus musculus
47	280	3.4	1459	1	CA12_MOUSE	P02453	rattus norv	103	233.5	2.8	815	1	MYG9_DROME	O93484	oncorhynch
48	278.5	3.4	1372	1	CA12_RAT	P28481	mus musculus	104	233.5	2.8	1003	1	ME06_HUMAN	O93484	oncorhynch
49	277.5	3.4	1460	1	CA11_CANFA	Q9x517	canis fami	105	233.5	2.8	1669	1	CA14_MOUSE	O93484	oncorhynch
50	277.5	3.4	1806	1	CA1B_HUMAN	P12107	homo sapien	106	232	2.8	1691	1	CA64_HUMAN	O93484	oncorhynch
51	277	3.4	1453	1	CA11_CHICK	P02457	gallus gall	107	231	2.8	754	1	5E5_RAT	O93484	oncorhynch
52	274	3.3	1838	1	CA15_HUMAN	P20908	homo sapien	108	228	2.8	825	1	CA54_CANFA	O93484	oncorhynch
53	273.5	3.3	1453	1	CA11_MOUSE	P11087	mus musculus	109	228	2.8	1466	1	IE18_PRYKA	O93484	oncorhynch
54	270	3.3	1464	1	CA11_HUMAN	P02455	homo sapien	110	227.5	2.8	460	1	RINI_HUMAN	O93484	oncorhynch
55	268.5	3.3	1464	1	CA21_RAT	P02466	rattus norv	111	225.5	2.7	1027	1	CARF_RIPRA	P33754	homo sapien
56	268.5	3.3	1603	1	CA1F_HUMAN	Q07092	homo sapien	112	225.5	2.7	2142	1	BAT2_HUMAN	P46634	homo sapien
57	268.5	3.3	1736	1	CA2B_MOUSE	Q64733	mus musculus	113	225	2.7	911	1	CA1B_BOVIN	Q28083	bos taurus
								114	225	2.7	1669	1	CA14_HUMAN	P02462	homo sapien

115	224	2.7	1403	1	B1RF_MOUSE	Q931b6 mus musculus
116	223	2.7	1527	1	CA1H_MOUSE	P39061 mus musculus
c 117	222.5	2.7	1446	1	IE18_PRIVA	P33479 pseudorabie
118	222	2.7	671	1	CA11_RAT	P02454 rattus norv
119	221.5	2.7	2321	1	NTC3_HUMAN	Q9um47 homo sapien
c 120	220.5	2.7	1461	1	IE18_PRIVT	P11675 pseudorabie
c 121	219.5	2.7	743	1	CA1B_MOUSE	O00760 mus musculus
c 122	219.5	2.7	744	1	CA1B_HUMAN	P27658 homo sapien
c 123	219.5	2.7	1712	1	CA2A_HUMAN	P08572 homo sapien
124	219	2.6	456	1	R1N1_RAT	P29315 rattus norv
125	218.5	2.6	911	1	CA1B_BOVIN	Q28083 bos taurus
126	218.5	2.6	2161	1	SHK1_HUMAN	Q9Y566 homo sapien
c 127	217.5	2.7	1516	1	CA1H_HUMAN	P39060 homo sapien
c 128	217.5	2.7	2161	1	SHK1_HUMAN	Q9Y566 homo sapien
c 129	216.5	2.6	867	1	SSPO_BOVIN	P98167 bos taurus
c 130	215.5	2.6	1402	1	B1RG_MOUSE	Q931b3 mus musculus
c 131	215	2.6	635	1	CA28_HUMAN	P25067 homo sapien
132	215	2.6	689	1	CA29_HUMAN	Q14055 homo sapien
133	214.5	2.6	635	1	CA29_HUMAN	P25067 homo sapien
134	214.5	2.6	1185	1	DRP1_HUMAN	P54259 homo sapien
c 135	214.5	2.6	1324	1	IRS2_HUMAN	Q9Y4h2 homo sapien
136	214	2.6	1756	1	CA14_CAEEL	P17139 caenorhabdi
137	213.5	2.6	1027	1	CA1F_RIFPA	P30754 riftia pach
c 138	212.5	2.6	921	1	CA19_HUMAN	P20949 homo sapien
c 139	212	2.6	825	1	ICP0_HSV2H	P28284 herpes simp
c 140	212	2.6	1691	1	CA64_HUMAN	Q14031 homo sapien
c 141	210	2.6	684	1	CA39_HUMAN	Q14050 homo sapien
c 142	209.5	2.5	684	1	CA39_HUMAN	Q14050 homo sapien
c 143	209.5	2.6	1336	1	W146_HUMAN	Q9c0f8 homo sapien
c 144	209.5	2.5	1403	1	B1R4_MOUSE	Q9c0f8 homo sapien
c 145	209.5	2.6	1756	1	CA14_CAEEL	P17139 caenorhabdi
146	209	2.5	5703	1	M03B_HUMAN	Q9h064 homo sapien
147	206.5	2.5	921	1	CA19_MOUSE	Q05722 mus musculi
c 148	205.5	2.5	2167	1	SHK1_RAT	Q9wv48 rattus norv
c 149	208	2.5	680	1	CA1A_HUMAN	O03692 homo sapien
c 150	207.5	2.5	1964	1	NTC4_MOUSE	P31695 mus musculi

ALIGNMENTS

RESULT 1  
 CARF\_HUMAN  
 ID CARF\_HUMAN STANDARD; PRT; 1040 AA.  
 AC Q9HC29; Q9GRH5; Q9GRH6; Q9GRH8;  
 DT 28-FEB-2003 (rel. 41, Created)  
 DT 28-FEB-2003 (rel. 41, Last sequence update)  
 DT 15-SEP-2003 (rel. 42, Last annotation update)  
 DE Caspase recruitment domain protein 15 (Nod2 protein) (Inflammatory  
 bowel disease protein 1).  
 GN CARD15 OR NOD2 OR IBD1.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBT\_taxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), MUTAGENESIS OF LYS-305, AND  
 RP VARIANT ARG-908.

RC TISSUE=Breast;  
 RA PubMed=11087742;  
 RA Ogura Y., Inohara N., Benito A., Chen F.F., Yamaoka S., Nunez G.;  
 RT "Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and  
 RT activates NF-kappaB."  
 RL J. Biol. Chem. 276:4812-4818(2001).  
 [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS CD THR-140; ARG-157;  
 RP CYS-235; ARG-248; SER-268; SER-289; ASN-291; VAL-301; TRP-311;  
 RP VAL-348; ARG-352; CYS-373; SER-414; LEU-431; VAL-432; LYS-441;  
 RP VAL-612; THR-612; TRP-684; TRP-702; CYS-703; CYS-713; GLY-725;  
 RP VAL-755; VAL-758; LYS-778; MET-793; LYS-843; SER-853; VAL-863;  
 RP ARG-908; ASP-918; ASP-924 AND ILE-955, VARIANTS ULCERATIVE COLITIS  
 RP THR-140 AND THR-885, AND VARIANT MET-189.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=21279172; PubMed=11389576;  
 RX Hugot J.-P., Chamaillard M., Zouali H., Lesage S., Cezard J.-P.,  
 RA Belaiche J., Almer S., Tykocil C., O'Morain C.A., Gassull M., Binder V.,  
 RA Finkel Y., Cortot A., Modigliani R., Laurent-Puig P.,  
 RA Gower-Rousseau C., Macry J., Colombel J.-F., Sabaatou M., Thomas G.,  
 RT "Association of NOD2 leucine-rich repeat variants with susceptibility  
 RT to Crohn's disease."  
 RL Nature 411:599-603(2001).  
 [3]  
 RN VARIANTS BS GLN-334; TRP-334 AND PHE-469.  
 RP MEDLINE=21419644; PubMed=11528384;  
 RX Miceli-Richard C., Lesage S., Rybojad M., Prieur A.M.,  
 RA Manouvrier-Hanu S., Harner R., Chamaillard M., Zouali H., Thomas G.,  
 RA Hugot J.-P.,  
 RT "CARD15 mutations in Blau syndrome."  
 RL Nat. Genet. 29:19-20(2001).  
 CC -1- FUNCTION: Induces NF-kappaB via RICK (CARDIAX, RIP2) and IKK-  
 CC gamma. Confers responsiveness to intracellular bacterial  
 CC lipopolysaccharides (LPS).  
 CC -1- SUBUNIT: Binds to RICK by CARD-CARD interaction.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative Initiation;  
 CC Comment=2 isoforms, 1/Nod2 (shown here) and 2/Nod2b, are  
 CC produced by alternative initiation. Both isoforms can activate  
 CC NF-kappaB. Isoform 1 is more abundant;  
 CC -1- TISSUE SPECIFICITY: Monocytes-specific.  
 CC [MIM:166580], a rare autosomal dominant disorder characterized by  
 CC early-onset granulomatous arthritis, uveitis and skin rash.  
 CC -1- DISEASE: Defects in CARD15 are a cause of susceptibility to  
 CC Crohn's disease (CD) [MIM:266600], a form of remitting  
 CC inflammatory bowel disease. CD may involve any part of the  
 CC gastrointestinal tract, but most frequently the terminal ileum and  
 CC colon. Bowel inflammation is transmural and discontinuous. Crohn's  
 CC disease is commonly classified as autoimmune disease.  
 CC -1- DISEASE: Defects in CARD15 are a cause of susceptibility to  
 CC ulcerative colitis [MIM:191390], a chronic inflammatory bowel  
 CC disease. In ulcerative colitis, the inflammation is continuous and  
 CC limited to rectal and colonic mucosal layers. Ulcerative colitis  
 CC is commonly classified as autoimmune disease.  
 CC -1- SIMILARITY: Contains 2 CARD domains.

```

CC -1- SIMILARITY: Contains 1 NACHT domain.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF178930; AAC3677.1; -.
DR EMBL; AF385089; AAK70867.1; -.
DR EMBL; AF385090; AAK70868.1; -.
DR EMBL; AJ303140; CAC42117.1; -.
DR Genew; HGNC:5331; CARD15.
DR MIM; 605956; -.
DR MIM; 186580; -.
DR MIM; 266600; -.
DR MIM; 191390; -.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR Pfam; PF00560; LRR_2.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PSS0209; CARD; 2.
DR PROSITE; PSS0837; NACHT; 1.
KW ATP-binding; Repeat; Leucine-rich repeat; Disease mutation;
KW Alternative initiation; Polymorphism.
FT CHAIN 1 1040
FT CASPASE RECRUITMENT DOMAIN PROTEIN 15,
FT ISOFORM 1.
FT CHAIN 28 1040
FT INIT MET 28 28
FT DOMAIN 26 122
FT DOMAIN 126 218
FT DOMAIN 293 618
FT NP_BIND 299 306
FT REPEAT 786 812
FT REPEAT 814 837
FT REPEAT 842 865
FT REPEAT 926 949
FT REPEAT 954 977
FT REPEAT 982 1005
FT VARIANT 140 140
FT VARIANT 157 157
FT VARIANT 189 189
FT VARIANT 235 235
FT VARIANT 248 248
FT VARIANT 268 268
FT /FTid=VAR_012670.

FT VARIANT 289 289
FT VARIANT 291 291
FT VARIANT 294 294
FT VARIANT 301 301
FT VARIANT 311 311
FT VARIANT 334 334
FT VARIANT 334 334
FT VARIANT 334 334
FT VARIANT 348 348
FT VARIANT 352 352
FT VARIANT 373 373
FT VARIANT 414 414
FT VARIANT 431 431
FT VARIANT 432 432
FT VARIANT 441 441
FT VARIANT 469 469
FT VARIANT 612 612
FT VARIANT 612 612
FT VARIANT 612 612
FT VARIANT 684 684
FT VARIANT 702 702
FT VARIANT 703 703
FT VARIANT 713 713
FT VARIANT 725 725
FT VARIANT 755 755
FT VARIANT 758 758
FT VARIANT 778 778
FT VARIANT 793 793
FT VARIANT 843 843
FT VARIANT 853 853
FT VARIANT 863 863

FT N -> S.
FT /FTid=VAR_012671.
FT D -> N (in CD).
FT /FTid=VAR_012672.
FT T -> S (in CD).
FT /FTid=VAR_012673.
FT A -> V (in CD).
FT /FTid=VAR_012674.
FT R -> W (in CD and ulcerative colitis).
FT /FTid=VAR_012675.
FT R -> Q (in BS).
FT /FTid=VAR_012676.
FT R -> W (in BS).
FT /FTid=VAR_012677.
FT L -> V (in CD).
FT /FTid=VAR_012678.
FT H -> R (in CD).
FT /FTid=VAR_012679.
FT R -> C (in CD).
FT /FTid=VAR_012680.
FT N -> S (in CD).
FT /FTid=VAR_012681.
FT S -> L (in CD).
FT /FTid=VAR_012682.
FT A -> V (in CD).
FT /FTid=VAR_012683.
FT E -> K (in CD).
FT /FTid=VAR_012684.
FT L -> F (in BS).
FT /FTid=VAR_012685.
FT A -> T (in CD).
FT /FTid=VAR_012686.
FT A -> V (in CD).
FT /FTid=VAR_012687.
FT R -> W (in CD).
FT /FTid=VAR_012688.
FT R -> W (risk factor for CD).
FT /FTid=VAR_012689.
FT R -> C (in CD and ulcerative colitis).
FT /FTid=VAR_012690.
FT R -> C (in CD).
FT /FTid=VAR_012691.
FT A -> G (in CD).
FT /FTid=VAR_012692.
FT A -> V (in CD and ulcerative colitis).
FT /FTid=VAR_012693.
FT A -> V (in CD).
FT /FTid=VAR_012694.
FT E -> K (in CD).
FT /FTid=VAR_012695.
FT V -> M (in CD).
FT /FTid=VAR_012696.
FT E -> K (in CD).
FT /FTid=VAR_012697.
FT N -> S (in CD).
FT /FTid=VAR_012698.
FT M -> V (in CD).

```

FT VARIANT 885 885 /FtId=VAR\_012699.  
A->T (in ulcerative colitis).  
Ft /FtId=VAR\_012700.  
FT VARIANT 908 908 G->R (in CD).  
Ft /FtId=VAR\_012701.  
FT VARIANT 918 918 A->D (risk factor for CD).  
Ft /FtId=VAR\_012702.  
FT VARIANT 924 924 G->D (in CD).  
Ft /FtId=VAR\_012703.  
FT VARIANT 955 955 V->I.  
Ft /FtId=VAR\_012704.  
FT MUTAGEN 305 K->R: NO ACTIVATION.  
SQ SEQUENCE 1040 AA; 115282 MW; 0037592D96D7DDEF CRC64;  
  
Alignment Scores:  
Pred. No.: 3,91e-295 Length: 1040  
Score: 5435.00 Matches: 1040  
Percent Similarity: 99.90% Conservative: 0  
Best Local Similarity: 99.90% Mismatches: 0  
Query Match: 65.72% Indels: 1  
DB: 1 Gaps: 0  
  
US-10-014-269-33 (1-4486) x CARF\_HUMAN (1-1040)  
  
QY 106 ATGGGAGAAAGGGGTGTTCAGCCTCTCACGATGAGAGAAAGAGCAATGTCTCTC 165  
Db 1 MetGlyGluGluGlySerAlaSerHisAspGluGluArgAlaSerValLeuLeu 20  
  
QY 166 GACATCTCCGGGTGTGAATGTCTCGCAGAGAGCTTTTCAGGCACAGAGCCAG 225  
Db 21 GlyHisSerProGlyCysGluMetCysSerGlnGluAlaPheGlnAlaGlnArgSerGln 40  
  
QY 226 CTGGTCAGAGCTGCTGGTCTCAGGGTCCTCGAGAGGCTTCAGAGTGCCTGGACTGGCTG 285  
Db 41 LeuValGluLeuLeuValSerGlySerLeuGluGluGlyPheGluSerValLeuAspTyrLeu 60  
  
QY 286 CTGTCTGGAGAGTCTCTCTCTGGAGAGACTACAGAGGCTTCACCTCCTGGAGCAGCCT 345  
Db 61 LeuSerTyrGluValLeuSerTyrGluAspTyrGluGlyPheHisLeuLeuGlyGlnPro 80  
  
QY 346 CTCTCCCACTTGGCAGGCGGCTCTGACACAGCGTCTGAAATAAAGGTACTGGCGCTGT 405  
Db 81 LeuSerHisLeuValAlaArgValLeuLeuAspThrValTyrPheAsnGlyGlyTyrTyrPheAsn 100  
  
QY 406 CAGAAAGCTCATCGGAGCTGCCCAAGAGGCCAGGCCAGACGCTGCCCAAGAGTGCAT 465  
Db 101 GlnLysLeuIleAlaAlaGlnGlnAlaGlnAlaAspSerGlnSerProLysLeuHis 120  
  
QY 466 GGCTGTGGAGCCCGCACTGCTCCACCAAGCCCGAGACCTGCAAGTCAACCGCCAGC 525  
Db 121 GlyCysTyrPheAspProHisSerLeuHisProAlaArgAspLeuGlnSerHisArgProAla 140  
  
QY 526 ATTGTCAAGAGGCTCCACAGGCAATGTGAGGAACATGCTGACCTGGCATGGAGAGCGGCT 585  
Db 141 IleValAlaArgValLeuHisSerHisValGlnAsnMetLeuAspLeuValTyrGlnGly 160  
  
QY 586 TTCGTACGCCAGTATGAAATGATGAAATCAGGTTGCCAATCTTCACACCGTCCCAAGG 645  
  
Db 161 PheValSerGlnTyrGluCysAspGluIleArgLeuProIlePheThrProSerGlnArg 180  
  
QY 646 GCAAGAGGCTGCTTATCTGCCACGCTGAAGACCAATGATGGCTGCTCTCTTA 705  
Db 181 AlaArgValLeuLeuAspLeuValAlaThrValLysAlaAsnGlyLeuValAlaPheLeuLeu 200  
  
QY 706 CAACATGTCAGGAATTAACAATCCCATGGCCCTGCTTGAAGACTGCCACATGCAAG 765  
Db 201 GlnHisValGlnGlnLeuProValPheLeuAlaLeuPheLeuGluAlaAlaThrCysLys 220  
  
QY 766 AAGTATAGCCCAAGTACAGACCAAGCTGTCTGCTCAGTCTGCTTCCATGACTAT 825  
Db 221 LysTyrMetAlaLysLeuArgThrThrValSerAlaGlnSerArgPheLeuSerThrTyr 240  
  
QY 826 GATGAGCAGACGACCTGCTGGAGACATATACAGAGAAATGCTCGGAGGCTG 885  
Db 241 AspGlyAlaGluThrLeuCysLeuGluAspIleTyrThrGlnAsnValLeuGluValTyr 260  
  
QY 886 GCAAGATGTGGCATGGCTGACCCCGCAGAGAGCCACCTGGGCTGAGAGAG 945  
Db 261 AlaAspValGlyMetAlaGlyProProGlnLysSerProAlaThrLeuGlyLeuGlu 280  
  
QY 946 CTCTTCAGCACCCCTGGCCACTCAATGACATCGGACACTGTGCTGGTGGGTAG 1005  
Db 281 LeuPheSerThrProGlyHisLeuAsnAspAspAlaAspThrValLeuValValGlyGln 300  
  
QY 1006 GCGGCAAGTGGCAAGACAGCTCCTGACAGGCGCTGCACTGTGTGGGCTGACAGG 1065  
Db 301 AlaGlySerGlyLysSerThrLeuLeuGlnArgLeuHisLeuLeuThrPheAlaGlyGln 320  
  
QY 1066 GACTTCAGGAATTTCTCTTTGCTTCCCATTCAGGCTCCGAGAGCTGCAGTGCATGGCC 1125  
Db 321 AspPheGlnGluThrLeuPheValAlaPheProPheSerCysArgGlnLeuGlnCysMetAla 340  
  
QY 1126 AAACACCTCTGTGCGGACTCTACTCTTGAACACTGTGTGGGCTGAGTGGTCA 1185  
Db 341 LysProLeuSerValArgThrLeuLeuPheGluHisCysCysTyrProAspValGlyGln 360  
  
QY 1186 GAAAGCATCTTCCAGTTACTCTCTTACCAACCTGACCGTGTCTGTAACCTTGAATGAC 1245  
Db 361 GluAspIlePheGlnLeuLeuLeuAspHisProAspArgValLeuLeuThrPheAspGly 380  
  
QY 1246 TTTCAGAGTCAAGTTCAGGTTCAAGGATCGTGAAGCCACCTGCTCCCAAGCCACCC 1305  
Db 381 PheAspGluPheLeuPheValPheThrAspArgGluArgHisCysSerProThrAspPro 400  
  
QY 1306 AACTGTGTCCAGACCTGCTCTTCAACCTTCTGACAGGCAACCTGCTGAAGATGCCGC 1365  
Db 401 ThrSerValGlnThrLeuLeuPheAsnLeuLeuGlnGlnAsnLeuLeuLysAsnAlaArg 420  
  
QY 1366 AAGTGGTGAACAGCGCTCCGCGGCTGTGTGGGCTTCTTCAAGAAATGATCGCACCC 1425  
Db 421 LysValValThrSerArgProAlaAlaValSerAlaPheLeuArgLysTyrIleArgThr 440  
  
QY 1426 GAGTTCACCTCAAGGCTTCTGAAACAGGCACTCAAGTGAATCGAAGAGGCGCAT 1485

Db	441	GIuPheAsnLeuLysGlyPheSerGluGlnGlyIleGluLeuTyrLeuArgLysArgHis	460
QY	1486	CATGAGCCCGGGGGTGGGAGACCGGCTCATCCGCTGCTCCAGAGACCTCAGCCCTGCAC	1345
Db	461	HisGluProGlyValAlaAspArgLeuIleArgLeuGlnGlnIleThrSerAlaLeuHis	480
QY	1546	GGTTTGGCCACCTGCTGCTCATGGATGGTGGCCAAATGGCCACCGAAGCTGTG	1605
Db	481	GlyLeuGlyHisLeuProValPheSerThrPheValSerLysGlyHisGlnGluLeu	500
QY	1606	CTGCAGAGGGGGGGTCCCAAAAGACACTACAGATAGTACTGCTGATTTCTGCACAT	1665
Db	501	LeuGlnGluGlySerProLysThrThrAspMetLysLeuLeuIleLeuGlnHis	520
QY	1666	TTTCTGCTGCATGCCACCCCGCCAGACTCAGCTTCCCAAGCTGGGAGCCGCTCTTT	1725
Db	521	PheLeuLeuHisAlaThrProProAspSerAlaSerGlnGlyLeuGlyProSerLeuLeu	540
QY	1726	CGGGGCGGCTCCCAACCTCCTGCACCTGGGCAAGTGGCTGTGGGGGCTGGGAGT	1785
Db	541	ArgGlyAcgLeuProThrIleLeuHisIleuGlyArgLeuAlaLeuIleProGlyLeuGlyMet	560
QY	1786	TGCTGCTACGTGTCTCAGCCCAAGACTCCAGGAGCAAGGTCAGCCCTGATGACATT	1845
Db	561	CysCysTyrValPheSerAlaGlnGlnLeuGlnAlaGlnValSerProAspAspIle	580
QY	1846	TCTCTTGGCTTCTGGTGGCTGCCAAAGGTGGTGGCGAGGAGTACGGGCGCCCTGAA	1905
Db	581	SerLeuGlyPheLeuValArgAlaLysGlyValValProGlySerThrAlaProLeuGlu	600
QY	1906	TTCCCTTCAATCACTTTCAGTGGCTTCTTGGCGGCTTCTACTGGGACATCAGTGTAT	1965
Db	601	PheLeuHisIleThrPheGlnCysPhePheAlaAlaPheTyrLeuAlaLeuSerAlaAsp	620
QY	1966	GTGCCACCAAGCTTGGTCAAGACCTCTTCAATTGGGAGGAGGCAACTCCCAATG	2025
Db	621	ValProProAlaLeuLeuArgHisIleuPheAsnProGlyArgProGlyAsnSerProMet	640
QY	2026	GCCAGGCTCCTGGCCACGATGTGCATCCAGGCTCGAAGGGAAGAGACAGAGGCTGGCA	2085
Db	641	AlaArgLeuLeuProThrMetCysIleGlnAlaSerGlnGlyLysAspSerValAla	660
QY	2086	GCTTGGTGCAGAAAGCCGAGCCGACACAACTTCATACATCAAGACCTCCTGGCAGGG	2145
Db	661	AlaLeuLeuGlnLysAlaGluProHisAsnLeuGlnIleThrAlaAlaPheLeuAlaGly	680
QY	2146	CTGTGTCCCGGGAGACTGGGGCTGCTGGCTGAGTGGCAGACATCTGAGAGAGCCCTG	2205
Db	681	LeuLeuSerArgGluHisIleTrpGlyLeuLeuAlaGluCysGlnHisSerGluLysAlaLeu	700
QY	2206	CTCCGGGCGCAGGCTGTGCCGCTGGTGTCTGGCCCGACACTCCGCAAGCACTTCCAC	2265
Db	701	LeuArgArgGlnAlaCysAlaArgTrpCysLeuAlaAspSerLeuArgLysPheHis	720
QY	2266	TCCATCCCGCAGCTGCACCGGGTGAAGGCGAAGAGGTGCATGCATGGCGGGTTGATC	2325
Db	721	SerIleProProAlaAlaProGlyGlnAlaLysSerValHisAlaMetProGlyPheIle	740
QY	2326	TGGCTCATCCGAGCCCTGTACAGATGCAGAGGAGGAGGCTGGCTGGAAAGCTGCAGCT	2385
Db	741	TrpLeuIleArgSerLeuTyrGluMetGlnGluLysAlaGlyLeuAlaArgLysAlaAlaArg	760
QY	2386	GGCTGAAATGTTGGGACCTCAAGTTGACATTTGGAGTGGGGCCCATGAGTGGCT	2445
Db	761	GlyLeuAsnValGlyHisLeuLysLeuThrPheCysSerValGlyProThrGlnCysAla	780
QY	2446	GGCTGGGCTTGTGCTGACAGACCTCCGGGGGCGCGGCGCTGCAGCTGGACTCAAC	2505
Db	781	AlaLeuAlaPheValLeuGlnHisLeuArgArgProValAlaLeuGlnLeuAspTyrAsn	800
QY	2506	TCTGGGGTGAATTGGGCTGAGAGAGTGGCTGGCTGGCTGGATGTCTGCAGAGGCTG	2565
Db	801	SerValGlyAspIleGlyValGlnGlnLeuLeuProCysLeuGlyValCysLysAlaLeu	820
QY	2566	TATTTGCCGATTAACATATCTGAGACGAGGAGTCTGCAGACTCATGAAATGCTCTT	2625
Db	821	TyrLeuArgAspAsnAsnIleSerAspArgGlyLysCysLysLeuIleGluCysAlaLeu	840
QY	2626	CAGTGCAGCAATTGCAGAAATTAGCTATTCACAAACAATTGACTGAGGCTGTGCA	2685
Db	841	HisCysGlnGlnLeuGlnLysLeuAlaLeuPheAsnAsnLysLeuThrAspGlySal	860
QY	2686	CAGTCAATGGCTAACTCTTGCATGCAGGCAAGACTTCTTGGCAATTGAGGCTGGGAAT	2745
Db	861	HisSerPheValLysLeuLeuAlaCysArgGlnAsnPheLeuAlaLeuArgLeuGlyAsn	880
QY	2746	AAGTACATCACTGGCGGAGGCCCAAGTCTGGCCGAGGGGCTCCGAGGCAACCTCC	2805
Db	881	AsnTyrIleThrAlaAlaGlyAlaGlnValLeuAlaGlnGlyLeuArgLysAsnThrSer	900
QY	2806	TTCAGTTTCTGGGATTCGTGGGCAACAAGTGGGTGACAGGGGGCCCAAGCCCTGGCT	2865
Db	901	LeuGlnPheLeuGlyPheTrpGlyAsnArgValGlyAspGlnGlyAlaGlnAlaLeuAla	920
QY	2866	GAAAGCTTGGGTGATCACAGGCTTGAGGTGGCTCAGCTGGTGGGGAACAACATTGGC	2925
Db	921	GluAlaLeuGlyAspHisGlnSerLeuArgTrpLeuSerLeuValGlyAsnAsnIleGly	940
QY	2926	AGTGGGATCCCAAGCCTTGGCACTGATGCTGGCAAGAACCTCATCTAGAAAGACTC	2985
Db	941	SerValGlyAlaGlnAlaLeuAlaLeuMetLeuAlaLysAsnValMetLeuGluGluLeu	960
QY	2986	TGCTGGAGGAGAACACTCTCCAGGATGAAGGTGTATGTTCTCTCCAGAAAGACTGAG	3045
Db	961	CysLeuGluGluLysAsnHisLeuGlnAspGlnGlyValCysSerLeuAlaGlnGlyLeuLys	980
QY	3046	AAAAATCAAGTTGAAAAATCCTGAAGTGTCCAAATAACTGCATCACCTACCTAGGGCA	3105
Db	981	LysAsnSerSerLeuLysIleLeuLysLeuSerAsnAsnCysAlaThrTyrLeuGlyAla	1000
QY	3106	GAAAGCCTCCTCGAGGCGCCCTTGAAGAAATACACCATCTGGAAGTCTGGCTCCGAGG	3165
Db	1001	GluAlaLeuLeuGlnAla-LeuGluArgAsnAspThrIleLeuGluValTrpLeuArgGln	1020



QY 3166 GACACTTCTCTAGAGAGTTGACAAAGCTCGGCTGACGGACACCAACTCTGCT 3225  
DB 1020 yAaRtHrPheserIeUaGluValAaPylSueGlyCySaRgSpThrArgLeuLeu 1040  
QY 3226 T 3226  
DB 1040 u 1040

RESULT 2  
CARD\_MOUSE  
ID CARD\_MOUSE STANDARD; PRT: 1020 AA.  
AC Q8K3Z0;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Caspase recruitment domain protein 15 (Nod2 protein).  
GN CARD15 OR NOD2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RP STRAIN=BALB/c; TISSUE=Monocytes;  
RA Iwagaya Y., Davey M.P., Martin T.M., Planck S.R., Depriest M.L.,  
RA Bugn M.M., Suing C., Rosenbaum J.T.;  
RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15  
RT gene."  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANTS ALA-212, ARG-240;  
RP CYS-422; VAL-485; ALA-603; ILE-675 AND GLN-925.  
RC STRAIN=NR1; TISSUE=Breast cancer;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stepleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,  
RA Brownstein W.J., Uceda T.B., Toshiyuki S., Carninci P., Plange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Meehan A., Rodigues S., Sanchez A.,  
RA Whiting M., Meehan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Scherich A., Schein J.E., Jones S.J.W., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: Induces NF-kappaB via RICK (CARD1X, RIP2) and IKK-  
CC gamma. Confers responsiveness to intercellular bacterial  
CC lipopolysaccharides (LPS) (By similarity).  
CC

CC -1- SUBUNIT: Binds to RICK by CARD-CARD interaction (By similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named Isoforms2;  
CC Name=1;  
CC IsoId=Q8K3Z0-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q8K3Z0-2; Sequence=VSP\_007069, VSP\_007070;  
CC Note=No experimental confirmation available;  
CC -1- SIMILARITY: Contains 2 CARD domains.  
CC -1- SIMILARITY: Contains 1 NACHT domain.  
CC -1- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EXBL: AF520774; AA076073.1; -  
DR EXBL: BC044774; AA044774.1; ALT\_INIT.  
DR WGD: WGI2429397; Card15.  
DR InterPro: IPR001315; CARD.  
DR InterPro: IPR007091; LRR\_NTHh.  
DR InterPro: IPR007111; NACHT\_NTPase.  
DR Pfam: PF00560; LRR\_2.  
DR PROSITE: PS02029; CARD; 2.  
DR PROSITE: PS00837; NACHT; 1.  
KW ATP-binding; Repeat; leucine-rich repeat; Polymorphism;  
KW Alternative splicing.  
FT DOMAIN 6 104 CARD 1.  
FT DOMAIN 106 200 CARD 2.  
FT DOMAIN 273 600 NACHT.  
FT REPEAT 685 709 LRR 1.  
FT REPEAT 726 749 LRR 2.  
FT REPEAT 766 792 LRR 3.  
FT REPEAT 794 817 LRR 4.  
FT REPEAT 822 845 LRR 5.  
FT REPEAT 850 873 LRR 6.  
FT REPEAT 882 906 LRR 7.  
FT REPEAT 906 929 LRR 8.  
FT REPEAT 934 962 LRR 9.  
FT REPEAT 963 985 LRR 10.  
FT REPEAT 1005 1019 LRR 9.  
FT VARSPLIC 1 7 Missing (in isoform 2).  
FT VARSPLIC 195 195 E -> EGVSLGSRSDNGFTLGLCFLL (in isoform  
FT VARSPLIC 2).  
FT VARSPLIC 212 212 /FTId=VSP\_007070.  
FT VARSPLIC 240 240 T -> A (in strain NR1).  
FT VARSPLIC 422 422 Q -> R (in strain NR1).  
FT VARSPLIC 485 485 L -> C (in strain NR1).  
FT VARSPLIC 485 485 G -> V (in strain NR1).  
FT VARSPLIC 603 603 V -> I (in strain NR1).  
FT VARSPLIC 675 675 V -> I (in strain NR1).  
FT VARSPLIC 925 925 E -> Q (in strain NR1).

SQ SEQUENCE 1020 AA; 113561 MW; 25504905 ECF70FB8 CRG64;

Alignment Scores: 8.73e-218 Length: 1020  
Pred. No.: 4183.00 Matches: 804  
Score: 4183.00 Conservation: 83  
Percent Similarity: 87.30% Mismatches: 128  
Best Local Similarity: 79.13% Indels: 1  
Query Match: 50.58% Gaps: 0

US-10-014-269-33 (1-4486) x CARF\_MOUSE (1-1020)

QY 181 TGTGAATGTGCTCCAGAGAGGCTTTTCAGGCAAGAGAGGAGCTGGTCCAGCTCTG 240  
DB 6 CysaspMetCysSerGlnGlnGlnPheGlnAlaGlnSerGlnLeuValAlaLeuLeu 25  
QY 241 GTCTCAGGGTCCCTGGAAGGCTTCAAGAGTCTGCTGCTGCTGCTGGAAGTTC 300  
DB 26 ILeSerGlySerLeuGlnGlnGlnGlnSerIleLeuAspTrpLeuLeuSerTrpVal 45  
QY 301 CTCTCCTGGAGGACTACAGAGGCTTCCACCTCCTGGAGGAGGCTCTCCACCTGGCC 360  
DB 46 LeuSerArgGlnAspTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 65  
QY 361 AGGCGCTCTTGAACCGCTCTGGAATTAAGGTACTTGGGCTGTCAAGAGCTCATCG 420  
DB 66 ArgArgLeuLeuAspThrValTrpAsnGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 85  
QY 421 GCTGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
DB 86 AlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 105  
QY 481 CACTGCTCCACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
DB 106 HisSerLeuHisProThrArgAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 125  
QY 541 CACAGCCATGTGGAGAGACATGCTGAGCTGGGATGGAGGAGGAGGAGGAGGAGGAGG 600  
DB 126 TyrAsnHisValGlnAlaMetLeuGlnLeuAlaArgGlnGlnGlnGlnGlnGlnGln 145  
QY 601 GAATGTGATGAATCAAGTTGCCATCTTCACACCGTCCAGAGGAGGAGGAGGAGGAGG 660  
DB 146 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 165  
QY 661 GATCTTGCACCGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
DB 166 AspLeuAlaAlaValLysAlaAsnGlyLeuAlaAlaPheLeuLeuGlnHisValArgGln 185  
QY 721 TTACAGATCCCATTTGGCTTGGAGGCTGCCACATGCAAGAGATATGGCAAG 780  
DB 186 LeuProAlaProLeuProLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 205  
QY 781 CTGAGGAGACAGGAGTCTGCTCAGTCTGCTTCTCACTATGATGAGGAGGAGGAGG 840  
DB 206 LeuArgThrMetValLeuThrGlnSerArgPheLeuSerThrTyrAspGlySerGln 225  
QY 841 CTCTGCTGAGGAGACATATACACAGAGATGTCTGAGAGGCTGGGAGATGTGGGAGT 900

DB 226 LeuGlyLeuGlnAspIleTyrThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 245  
QY 901 GCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
DB 246 AlaGlnGlnAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 265  
QY 961 GGCACCTCAATGAGGATGCCAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 266 GlyHisLeuLeuAsnArgAspAlaAspThrIleLeuValValGlnGlnGlnGlnGlnGln 285  
QY 1021 AGCAGGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080  
DB 286 SerThrLeuLeuGlnArgLeuHisLeuLeuTrpAlaTrpGlnArgSerPheGlnGlnPhe 305  
QY 1081 CTCTTTGCTTCCCATTCAGTACCTGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
DB 306 LeuPheIlePheProPheSerGlyArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 325  
QY 1141 CGAGCTCACTCTTGAACAGCTGCTGTTGGCTGATGTTGGTCAAGAGACATCTTCAAG 1200  
DB 326 ArgThrLeuLeuPheGlnHisCysCysTrpProAspValAlaGlnAspAspValPheGln 345  
QY 1201 TTACTCTTGAACACCTGACCGTGTCTGTAACTTTGATGGCTTGGAGAGTTCAAG 1260  
DB 346 PheLeuLeuAspHisProAspArgValLeuLeuThrPheAspGlyLeuAspGlnPheGln 365  
QY 1261 TTCAAGTTCAAGGATCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
DB 366 PheArgPheThrAspArgGlnArgHisAspSerProIleAspProThrSerValGlnThr 385  
QY 1321 CTGCTTTCAACCTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380  
DB 386 LeuLeuPheAsnLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 405  
QY 1381 CGTCCGCGCTGTGTGTGCTCTCAAGAGATCATCCGACGAGGAGGAGGAGGAGGAGG 1440  
DB 406 ArgProAspAlaValSerAlaLeuLeuArgGlyPheValArgThrGlnLeuGlnLeuGln 425  
QY 1441 GGCTTCTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500  
DB 426 GlyPheSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 445  
QY 1501 GCGAGCGCTCATCCGCTGCTCCAGAGACCTCAGGCTGAGGAGGAGGAGGAGGAGG 1560  
DB 446 AlaAspArgLeuIleGlnLeuIleGlnAlaThrSerAlaLeuHisGlyLeuGlnGln 465  
QY 1561 CCTGTCTTCATGATGCTGTCCAAATGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620  
DB 466 ProValPheSerTrpMetValSerArgCysHisArgGlnLeuLeuGlnGlnGlnGln 485  
QY 1621 TCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680  
DB 486 PheProThrThrSerThrAspMetCylLeuLeuIleLeuGlnHisPheLeuLeuHisAla 505  
QY 1681 ACCGCCCAAGATCAGCTTCCAGAGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740

Db 506 SerProAspSerSerProLeuGlyProGlyLeuLeuGlnSerArgLeuSer 525  
 QY 1741 ACCCTCTGCACTGGGCAAGACTGGCTCTGTGGGGCCCTGGGCAATGCTGCTAGCTGTTG 1800  
 Db 526 ThrLeuLeuHisLeuGlyHisLeuAlaLeuAlaGlyLeuAlaLeuSerCysTyrValPhe 545  
 QY 1801 TCAGCCCAAGCACTCCAGGCAAGCAAGCTCAAGCCCTGATGACATTTCTGTGGCTGCTG 1860  
 Db 546 SerAlaGlnGlnLeuGlnAlaGlnValAspSerAspAspIleSerLeuGlyPheLeu 565  
 QY 1861 GTGGCTGCCAAAGGTGCTGCTGCAAGGAGTACGGCCGCTTGAATTCCTTCACTGACT 1920  
 Db 566 ValArgAlaGlnSerSerValProGlySerTyrAlaProLeuGlnPheLeuHisIleThr 585  
 QY 1921 TTCAGTGTCTCTTGGCCGCTCTACCTGGGCACTCAAGTGGCTGATGGCAACCACTTTG 1980  
 Db 586 PheGlnCysPhePheAlaLeuPheTyrLeuAlaValSerAlaAspHisSerValAlaSer 605  
 QY 1981 CTCAAGACACTCTTCAATTGTGGCAAGGCAAGCACTCAACATGGCCAGGCTGCTGCC 2040  
 Db 606 LeuTyrHisLeuPheSerCysGlyArgLeuGlySerSerLeuLeuGlyArgLeuLeuPro 625  
 QY 2041 ACGATGTGCATCCAGGCTCCAGGCGGGAAGAAGCAAGCAAGCTTGTGCTGCAAG 2100  
 Db 626 AsnLeuCysIleGlnGlySerArgValLysGlySerGlnValAlaLeuLeuGlnLys 645  
 QY 2101 GCCAGGCCGCAACAACCTTCAATCAAGCAAGCTTCTTGGCAAGGCTGTGCTCCGGAAG 2160  
 Db 646 AlaGlnProHisAsnLeuGlnIleThrAlaAlaPheLeuAlaGlyLeuLeuSerGlnGln 665  
 QY 2161 CACTGGGCGCTGCTGCTGATGCTGCAAGCACTGAGAAAGGCGCTGCTCCGGCGCAAGCC 2220  
 Db 666 HisArgAspLeuLeuAlaLeuCysGlnValSerGlnArgValLeuLeuGlnArgGlnAla 685  
 QY 2221 TGTGCCGCTGCTGCTGCTGGCCGCGAGCTCCGCAAGCACTTCCACTCCATCCGCGAGCT 2280  
 Db 686 ArgAlaArgSerCysAlaLeuAlaHisSerLeuLysGlnHisPheHisSerIleProPheAla 705  
 QY 2281 GCACGGGTGAGGCAAGGCTGCTGCAAGCACTGCGCGGGTTCATCTGCTGATCCGAGC 2340  
 Db 706 ValProGlyGlnThrLysSerMetHisAlaSerProGlyPheIleTrpLeuIleArgSer 725  
 QY 2341 CTGTACAGATGCAAGGAGGAGGCGGCTGGTGGAGAGGCTGCAAGTGGCTGATGATG 2400  
 Db 726 LeuTyrGlnLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 745  
 QY 2401 CACTCAAGTTGACATTTGCACTGTGGAGTGGCCCACTGAGTGTGCTGCTGCTGCTTGTG 2460  
 Db 746 HisLeuLysLeuThrPheCysArgValGlyProAlaGlnCysAlaAlaLeuAlaPheVal 765  
 QY 2461 CTGCAGCAAGCTCGGGGCGGCGCTGGCCCTGAGCTGAGCTACAAGCTGTGTGGTGA 2520  
 Db 766 LeuGlnHisLeuGlnArgProValAlaLeuGlnLeuAspTyrAsnSerValGlyAspVal 785  
 QY 2521 GAGCTGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580  
 Db 786 GlyValGlnGlnLeuArgProCysLeuGlyValCysThrAlaLeuLysLeuArgAspAsn 805

QY 2581 AATATCTAGACCGAGGCATCTGCAAGCTCATTAATGATGCTCTTCACTGGGCAAGT 2640  
 Db 806 AsnIleSerAspArgGlyAlaArgThrLeuValGlnCysAlaLeuLysGlyGlnGlnLeu 825  
 QY 2641 CAAGAATAGCTCATTCACACAAATTAAGTGAAGGCTGTGCACTCCATGCGTAAG 2700  
 Db 826 GlnLysLeuAlaLeuPheAsnAsnLysLeuThrAspAlaCysAlaCysSerMetValLys 845  
 QY 2701 CTCTTGCATGACGAGCAAGCTTCTGGCAATGAGGCTGGGGAATTAATCACTGCTG 2760  
 Db 846 LeuLeuAlaHisLysGlnAsnPheLeuSerLeuLysGlyValGlnAsnHisIleThrAla 865  
 QY 2761 GCGGAGCCCAAGTGTGCTGCGCAAGGCGCTCCAGAGGCAACACTCTTGCATCTCTG 2820  
 Db 866 AlaGlyAlaGlnValLeuAlaGlnGlnLysSerAsnThrSerLeuLysPheLeuGly 885  
 QY 2821 TTCTGGGGCAAGAGTGGTGAAGAGGGGCGGCGGCTTGGCTGAGCCTTGGGCTGAT 2880  
 Db 886 PheTrpGlyAsnSerValGlyAspLysGlyThrGlnAlaLeuAlaGlnValAlaAsp 905  
 QY 2881 CACCAAGACTTGAAGTGGCTCAAGCTGGTGGGGAACAATTGGCACTGTGGTGGCCA 2940  
 Db 906 HisGlnAsnLeuLysTrpLeuSerLeuValGlyAsnAsnIleGlySerMetGlyAlaGln 925  
 QY 2941 GCCTTGGGCACTGATGCTGGCAAGAAAGCTCATCTAGAAAGACTTGGCTGAGAGAAC 3000  
 Db 926 AlaLeuAlaLeuMetLeuGlnGlnLysAsnLysSerLeuGlnGlnLysCysLeuGlnGln 945  
 QY 3001 CATCTCAAGATGAAGGTGATGTTCTCTGCAAGAGCTGAAAGAAATTCATGATTG 3060  
 Db 946 HisIleCysAspGlnGlyValTyrSerLeuAlaGlnGlnLysLysArgAsnSerThrLeu 965  
 QY 3061 AAAATCTGAAGTGTCCAAATACTGCACTCACTAGGCGCAAGAGCCCTCTCGAG 3120  
 Db 966 LysPheLeuLysLeuSerAsnAsnGlyIleThrTyrArgGlyAlaGlnAlaLeuLeuGln 985  
 QY 3121 GCCCTTGAAGGAATGACACCACTCTGGAAGTGGCTGCGAGGGAACACTTCTCT 3180  
 Db 986 AlaLeuSerArgAsnSerAlaIleLeuGlnValTrpLeuLysGlyAsnThrPheSerLe 1005  
 QY 3181 AAGAGAGTTGACAAAGCTCGGCTGCAAGGACCAAGCACTTGTCTT 3226  
 Db 1005 GlnGlnLysGlnThrLeuSerSerArgAspAlaArgLeuLeu 1020

Search completed: January 7, 2004, 19:48:56  
 Job time : 435.754 secs

OM nucleic - protein search, using frame\_plus\_nzp model

Run on: January 7, 2004, 19:25:21 ; Search time 170.769 Seconds

(Without alignments)  
1357.774 Million cell updates/sec

Title: US-10-014-269-33

Perfect score: 8270  
Sequence: 1 gtagacagatccagcgtccac.....ataactgttggtcaaac 4486

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 630525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=frame+\_nzp.model -DEV=xlp

Q=/usr2\_1/USPTO\_spool\_P/US10014269/runat\_07012004\_175229\_24521/app\_query.fasta\_1

-9358

-DAS=SPRMBL\_23 -QFMT=fastn -SUFFIX=rspt -MINMATCH=0.1 -IOPOOL=0 -IOPEXT=0

-DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL

-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10014269 @CGN 1.1 482 @runat 07012004 175229 24521 -NCPU=6 -ICPU=3

-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -IONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6

-FGAPEXT=7 -XGAPOP=10 -FGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organism:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacterioph:\*

17: sp\_archaeoph:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4183	50.6	1020	11 Q8K320	Q8K320 mus musculus
2	1209.5	14.6	953	11 Q8BH80	Q8BH80 mus musculus
3	1196.5	14.5	953	4 Q8WTF3	Q8WTF3 homo sapien
4	1106	12.2	706	11 Q8BU76	Q8BU76 mus musculus
5	495.5	6.0	1033	11 Q8R4B8	Q8R4B8 mus musculus
6	483	5.8	892	4 Q8NEU4	Q8NEU4 mus musculus
7	431.5	5.2	1052	7 Q8G3D8	Q8G3D8 ratu mus norv
8	431.5	5.2	1073	7 Q8G3D9	Q8G3D9 ratu mus norv
9	431.5	5.2	1153	7 Q8G3E0	Q8G3E0 ratu mus norv
10	417.5	5.0	1155	7 Q8P2P1	Q8P2P1 ratu mus norv
11	410.5	5.0	778	4 Q8NF48	Q8NF48 homo sapien
12	380	4.6	713	6 Q95L27	Q95L27 macaca fasc
13	357.5	4.3	733	4 Q8TEE2	Q8TEE2 mus musculus
14	354.5	4.3	982	11 Q8BU40	Q8BU40 mus musculus
15	352.5	4.3	499	4 Q8TEI3	Q8TEI3 homo sapien
16	352	4.3	5146	6 Q8SPM4	Q8SPM4 bos taurus
17	349.5	4.2	5146	6 Q8SPM4	Q8SPM4 bos taurus
18	339.5	4.1	1056	4 Q8NF42	Q8NF42 homo sapien
19	332	4.0	846	4 Q8IXT0	Q8IXT0 homo sapien
20	322	3.9	932	4 Q8FKL4	Q8FKL4 homo sapien
21	321.5	3.9	660	4 Q8NF06	Q8NF06 homo sapien
22	321	3.9	692	4 Q8GD51	Q8GD51 homo sapien
23	320	3.9	1464	11 Q8BIM4	Q8BIM4 mus musculus
24	319	3.9	884	7 Q2S675	Q2S675 mus musculus
25	317	3.9	1222	11 Q8K173	Q8K173 mus musculus
26	316.5	3.8	1787	10 Q9M4X9	Q9M4X9 chlamydomon
27	316	3.9	4123	4 Q7S851	Q7S851 homo sapien
28	313	3.8	1453	11 Q63079	Q63079 ratu mus norv
29	312.5	3.8	1464	4 Q8R473	Q8R473 mus musculus
30	310.5	3.8	1637	6 Q9XSV8	Q9XSV8 bos taurus
31	309	3.8	1464	11 Q8BK12	Q8BK12 mus musculus
32	308	3.8	1461	4 Q76045	Q76045 mus musculus
33	307.5	3.7	977	7 Q8HM99	Q8HM99 mus musculus
34	307.5	3.7	1222	11 Q8K173	Q8K173 mus musculus
35	306.5	3.7	863	11 Q8C6V9	Q8C6V9 mus musculus
36	306.5	3.7	1445	13 Q93251	Q93251 rana catesbe
37	304.5	3.7	1464	11 Q8BIM4	Q8BIM4 mus musculus
38	303.5	3.7	1314	11 Q8CGA5	Q8CGA5 mus musculus
39	303.5	3.7	1347	4 Q96Q83	Q96Q83 homo sapien
40	303	3.7	774	12 Q41971	Q41971 murid heppe

c	41	301	3.7	1118	5	Q9v13	Q9v13 drosophila
c	42	300.5	3.6	1464	11	Q8BKr2	Q8BKr2 mus musculus
c	43	300.5	3.7	1953	5	Q8B1T7	Q8B1T7 nephila mad
c	44	300	3.7	1450	13	Q9y1B4	Q9y1B4 cynops pyr
c	45	299.5	3.7	1160	4	Q14046	Q14046 homo sapien
c	46	299	3.6	1447	13	Q91B91	Q91B91 xenopus lae
c	47	299	3.6	1487	6	Q77753	Q77753 canis faml
c	48	298.5	3.6	1487	4	Q14047	Q14047 homo sapien
c	49	296	3.6	2703	5	Q9vE67	Q9vE67 drosophila
c	50	294.5	3.6	1418	6	Q28386	Q28386 equus caball
c	51	294.5	3.6	2944	11	Q63870	Q63870 mus musculus
c	52	292.5	3.6	1682	11	Q9CZ89	Q9CZ89 mus musculus
c	53	292	3.5	397	11	Q8BJF4	Q8BJF4 mus musculus
c	54	292	3.5	886	4	Q9NUS7	Q9NUS7 homo sapien
c	55	292	3.6	1420	13	Q90W37	Q90W37 gallus gall
c	56	291.5	3.6	890	12	Q8UZE1	Q8UZE1 ceratopithec
c	57	291	3.5	1414	5	Q26634	Q26634 strongyloce
c	58	290	3.5	1580	11	Q8BLX7	Q8BLX7 mus musculus
c	59	289.5	3.5	1366	4	Q15177	Q15177 homo sapien
c	60	289.5	3.5	1442	11	Q62033	Q62033 mus musculus
c	61	289.5	3.5	2715	5	Q61603	Q61603 drosophila
c	62	289.5	3.5	2716	5	Q81N94	Q81N94 mus musculus
c	63	289.5	3.5	4998	11	Q8C865	Q8C865 mus musculus
c	64	289	3.5	1637	6	Q8XSV8	Q8XSV8 bos taurus
c	65	288.5	3.5	1487	11	Q61431	Q61431 mus musculus
c	66	288	3.5	1097	4	Q9H6V0	Q9H6V0 homo sapien
c	67	287.5	3.5	706	12	Q41972	Q41972 murid herpe
c	68	287.5	3.5	4123	4	Q75851	Q75851 homo sapien
c	69	284.5	3.5	4998	11	Q8C865	Q8C865 mus musculus
c	70	284	3.5	1163	4	Q8N6U4	Q8N6U4 homo sapien
c	71	284	3.5	1419	11	Q63123	Q63123 rattus norv
c	72	284	3.5	1459	11	Q62032	Q62032 mus musculus
c	73	282.5	3.4	1442	11	Q62031	Q62031 mus musculus
c	74	282	3.4	312	11	Q8BV65	Q8BV65 mus musculus
c	75	282	3.4	1458	13	Q91089	Q91089 oncohychnu
c	76	282	3.4	1669	11	Q9QZ50	Q9QZ50 mus musculus
c	77	282	3.4	1802	5	Q17163	Q17163 brugia mala
c	78	281.5	3.4	1626	4	Q8NFM1	Q8NFM1 homo sapien
c	79	281.5	3.4	1684	6	Q8H1C1	Q8H1C1 canis faml
c	80	281	3.4	1414	5	Q26634	Q26634 strongyloce
c	81	280	3.4	1838	4	Q15094	Q15094 homo sapien
c	82	279	3.4	1420	13	Q90W37	Q90W37 gallus gall
c	83	278	3.4	727	12	Q41973	Q41973 murid herpe
c	84	277.5	3.4	1450	13	Q9Y1B4	Q9Y1B4 cynops pyr
c	85	276.5	3.3	1183	4	Q8N6U4	Q8N6U4 homo sapien
c	86	276	3.3	1745	4	Q9N206	Q9N206 homo sapien
c	87	276	3.3	1840	11	Q9J103	Q9J103 rattus norv
c	88	275	3.3	1462	11	Q9CZ89	Q9CZ89 mus musculus
c	89	274	3.3	1449	13	Q910C0	Q910C0 oncohychnu
c	90	274	3.3	1449	13	Q910C0	Q910C0 oncohychnu
c	91	273	3.3	1461	4	Q76045	Q76045 homo sapien
c	92	273	3.3	1464	4	Q8N473	Q8N473 homo sapien
c	93	272.5	3.3	1689	11	Q9QZ50	Q9QZ50 mus musculus
c	94	272.5	3.3	1691	11	Q9P5Q2	Q9P5Q2 mus musculus
c	95	271	3.3	223	4	Q8N102	Q8N102 homo sapien
c	96	270.5	3.3	1487	6	Q77753	Q77753 canis faml
c	97	269.5	3.3	608	12	Q9Q5K9	Q9Q5K9 herpesvirus
c	98	269.5	3.3	1418	6	Q28396	Q28396 equus caball
c	99	268.5	3.2	605	10	Q93ZV8	Q93ZV8 arabidopsis
c	100	268.5	3.3	1838	11	Q88207	Q88207 mus musculus
c	101	268	3.2	1998	11	Q8CFM4	Q8CFM4 mus musculus
c	102	268	3.2	1442	11	Q62031	Q62031 mus musculus
c	103	268	3.2	1442	11	Q62033	Q62033 mus musculus
c	104	268	3.2	1439	11	Q62032	Q62032 mus musculus
c	105	267.5	3.2	1626	4	Q8NFM1	Q8NFM1 homo sapien
c	106	267.5	3.2	1820	13	Q91B91	Q91B91 xenopus lae
c	107	267	3.2	1447	13	Q91B91	Q91B91 xenopus lae
c	108	266.5	3.2	1439	5	Q97406	Q97406 halictis di
c	109	266	3.2	1378	5	Q97405	Q97405 halictis di
c	110	266	3.2	1453	11	Q63079	Q63079 rattus norv
c	111	266	3.2	1621	4	Q9H4R9	Q9H4R9 homo sapien
c	112	265.5	3.2	1487	4	Q14047	Q14047 homo sapien
c	113	265.5	3.2	1739	11	Q9J112	Q9J112 mus musculus
c	114	264.5	3.2	1382	13	Q90V70	Q90V70 brachydanto
c	115	264.5	3.2	1835	13	Q91AU4	Q91AU4 gallus gall
c	116	264	3.2	1840	11	Q60467	Q60467 cricetus
c	117	263.5	3.2	1481	13	Q91718	Q91718 xenopus lae
c	118	263	3.2	1894	5	Q8MWS4	Q8MWS4 mytilus gal
c	119	262.5	3.2	1160	4	Q14046	Q14046 homo sapien
c	120	262	3.2	195	4	Q8N101	Q8N101 homo sapien
c	121	262	3.2	1737	11	Q9J104	Q9J104 rattus norv
c	122	262	3.2	1827	13	Q8U0N5	Q8U0N5 cryzias lat
c	123	261.5	3.2	1684	6	Q8H1C1	Q8H1C1 canis faml
c	124	261	3.2	1419	11	Q63123	Q63123 rattus norv
c	125	260.5	3.1	1445	13	Q93Z51	Q93Z51 rana catesb
c	126	260	3.1	166	4	Q15177	Q15177 homo sapien
c	127	259	3.1	620	16	Q8XZM9	Q8XZM9 ralstonia s
c	128	259	3.2	998	11	Q8CFM4	Q8CFM4 mus musculus
c	129	259	3.2	1835	13	Q91AU4	Q91AU4 gallus gall
c	130	258.5	3.1	1739	11	Q8U112	Q8U112 mus musculus
c	131	258	3.1	1449	13	Q8AM17	Q8AM17 brachydanto
c	132	258	3.1	1838	11	Q88207	Q88207 mus musculus
c	133	257.5	3.1	2703	5	Q9VE67	Q9VE67 drosophila
c	134	257.5	3.1	2715	5	Q61603	Q61603 drosophila
c	135	257.5	3.1	2716	5	Q81N94	Q81N94 drosophila
c	136	257	3.1	1258	13	Q8AM11	Q8AM11 brachydanto
c	137	257	3.1	1314	11	Q8C8K5	Q8C8K5 mus musculus
c	138	256.5	3.1	926	5	Q9W3G1	Q9W3G1 drosophila
c	139	256.5	3.1	1691	11	Q9P5Q2	Q9P5Q2 mus musculus
c	140	256.5	3.1	1838	4	Q15094	Q15094 homo sapien
c	141	255.5	3.1	868	13	Q90796	Q90796 gallus gall
c	142	255.5	3.1	1418	13	Q9W7R9	Q9W7R9 cynops pyr
c	143	255.5	3.1	1491	13	Q91718	Q91718 xenopus lae
c	144	254	3.1	1840	11	Q90103	Q90103 rattus norv
c	145	254	3.1	966	6	Q93UD1	Q93UD1 sus scrofa
c	146	252.5	3.1	1439	5	Q96C06	Q97406 halictis di
c	147	252.5	3.1	2157	4	Q96C06	Q96C06 mus musculus
c	148	252	3.1	1486	13	Q91717	Q91717 xenopus lae
c	149	252	3.0	1840	11	Q60467	Q60467 cricetus

ALIGNMENTS

```

RESULT 1
SEQ#20
AC      ORK320          PRELIMINARY;       PRT;   1020 AA.
DT      01-OCT-2002 (T-EMBLrel. 22, Created)
DT      01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DE      01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE      CARD15.
GN      CARD15.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;
RN      NM_
RP      SEQUENCE FROM N.A.
RC      STRAIN=BALB/c;
RA      Imanaga Y., Davey M.P., Martin T.M., Planck S.R., DePiret M.L.,
RA      Baugh M.M., Sung C., Rosenbaum J.T.;
RT      "Cloning, sequencing and expression analysis of the murine Nodl/Cards
RT      gene."
RL      Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RS      EMBL: AF520774; AAA76073.1; -.
DR      MGd; MG1:2429387; Card15.
DR      InterPro; IPR001315; CARD.
DR      InterPro; IPR007091; LRR_RNLMn.
DR      InterPro; IPR007111; NACT_NTPase.
DR      InterPro; IPR001130; TatD_DNase.
DR      PROSITE; PSS0209; CARD_2.
DR      PROSITE; PSS0503; LRR_Rf_1.
DR      PROSITE; PSS0837; NACTr_1.
DR      PROSITE; PSS10137; TATD_1.
DR      PROSITE; PSS10137; TATD_1.
SQ      SEQUENCE   1020 AA;  113561 MW;  25504905ECCF70FB6 CRC64;

Alignment Scores:
Pred. No.:           1,15e-289              Length:         1020
Score:               4183.00                 Matches:         804
Percent Similarity:  87.30%                  Conservative:    83
Best Local Similarity: 79.13%                Mismatches:     128
Query Match:        50,58%                   Indels:          1
Db:                 11                       Gaps:            0

US-10-014-269-33 (1-4486) x ORK320 (1-1020)
QY      181 TGTTAAATGTGTCGCAGAGAGCTTTTCAGGCACAGAGCAGCGTGGTCAACTGCG 240
Dc      6 CysApMwtCysSeedInclngUgUpRegInAlgnAgsSerGlnEuVatAlalaLeuIou 25
QY      241 GTCTCAGGGGTCCTGAAAAGGCTTCAGAGTGCTGCTGAAGTGGACTGGCTGCTG 300
Db      26 TTSeerGIglyserIenouclngUgPhoeIgUsertLeuVaSPTrpleuSenstRPApyVal 45
QY      301 CTCTCCTGGAGGAGACTACGAGGGAGCTTCACCTCGTGGGCGAGCGCTCTGCCACTTGGC 360
Dc      46 LeuSerArgIuApSfyGluGlyLeuSerLeuPregIInProleuSerHisSerAla 65
QY      361 AGGCGCCCTCTCGACACCCCTGTGAAATAAGGGTAAGTGGGCGCTGTGAAGAAGCTCATGCG 420

```

QY 1261 TTGAGGTTCAGGATGTCAGACGCCACTGCTCCCAAGCCACCTCTGTCCAGACC 1320  
Db PheArgPheThrAspArgGluArgHisCysSerProIleAspProIleSerValGlnThr 385  
QY 1321 CTGCTCTTCAACCTTCTGAGGGGCAACCTGCTGAGAGATGCCCGAProIleSerValGlnThr 1380  
Db LeuLeuPheAsnLeuLeuGlnGlyAsnLeuLeuLeuLysAsnHisAlaCysIleValLeuThrSer 405  
QY 1381 CTCGCGGCGCTGTGTGGCGCTTCTCAAGAAATCAATCCGACCGAATTCAACCTCAAG 1440  
Db ArgProAspAlaValSerAlaLeuLeuArgLysPheValArgThrGlnLeuGlnLeuLys 425  
QY 1441 GGCTCTCTGAGACGGGATCAGACTGTACCTGAGAGAACCGCACTCATGAGCGCGGGGTG 1500  
Db GlyPheSerGlnGlnGlyIleGlnLeuLysLeuLysLysHisSerGlnLysProGlyVal 445  
QY 1501 GCGAGCGGCTCATCCGCTCTCTCAAGAGACTCAGCGCTGCAAGGTTTGTGCACTTG 1560  
Db AlaAspArgLeuIleGlnLeuIleGlnAlaThrSerAlaLeuHisGlyLeuCysHisLeu 465  
QY 1561 CCTGCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620  
Db ProValPheSerArgPheValSerArgCysHisSerGlnLeuLeuLeuLeuLeuLeuLeu 485  
QY 1621 TCCCAAAAGCACTCAAGATATGATCTGCTGATCTGCAAGATTTTCTGCTGATGCC 1680  
Db PheProThrThrThrThrAspMetLysLeuLeuIleLeuGlnHisPheLeuLeuHisAla 505  
QY 1681 ACGCGCGCGCACTCAGCTTCCCAAGGCTGCGGAGCGCAAGCTTCTTGGGGCGCGCTCCC 1740  
Db SerProProAspSerSerProLeuGlnLeuGlyProGlyLeuLeuGlnSerAlaGlnLeuSer 525  
QY 1741 ACGCTCTGACCTTGGGCAAGACTGAGCTGTGTGGGGCTTGGGCAATGCTGCTACGTGTTG 1800  
Db ThrLeuLeuHisLeuGlnLysHisLeuAlaLeuLysGlyLeuAlaMetSerCysLysValPhe 545  
QY 1801 TCAGCGCGCAAGCTTCCAGGCAAGCTCAGGCTGATGACATTTCTGCTGAGCTTCTG 1860  
Db SerAlaGlnGlnLeuGlnAlaAlaGlnValAspSerAspAspIleSerLeuGlnLysPheLeu 565  
QY 1861 GTGCGTGCAGAAAGTGTGTGCCAGGGAATACGGCGCGCTGAGATTCCTTCAACACT 1920  
Db ValArgAlaGlnSerSerValProGlySerLysAlaProLeuGlnPheLeuHisIleThr 585  
QY 1921 TTCCAGTGTCTTGTGGCGGCTTCAACTGCACTCAGTGTGATGTGCAACAGCTTTG 1980  
Db PheGlnCysPhePheAlaAlaPheTyrLeuAlaValSerAlaAspThrSerValAlaSer 605  
QY 1981 CTCAGACACTCTTCATATTGTGGCAGGCGCAAGCTCAGCAATGGCGAGGCTCTGCC 2040  
Db LeuLysHisLeuPheSerCysGlyLysArgLeuLysSerSerLeuLeuAlaValLeuLeuPro 625  
QY 2041 ACGATGTGATCCAGGCGCTTGGAGGGAAGCAAGCAAGCTGCAAGCTTGTGCAAG 2100  
Db AsnLeuCysIleGlnGlySerArgValLysLysGlySerGlnAlaAlaLeuLeuGlnLys 645

QY 2101 GCGAGCGCGCAACACTTTCAGATCAAGACAGCTTCTGAGGAGGCTGTGTCCCGGAG 2160  
Db AlaGlnProHisAsnLeuGlnIleThrAlaAlaPheLeuAlaGlyLeuLeuSerGlnGln 665  
QY 2161 CAGCTGAGGCGCTGTGGCTGAGTGGCGCAATCTGAGGAGCGCTGCTCCGGCGCGAGCC 2220  
Db HisAlaGlnPheLeuAlaAlaCysGlnValSerGlnLysGlyAlaLeuLeuGlnArgAla 685  
QY 2221 TGTGCGCGCTGTGTGTGGCGCGAGCTCCGCAAGCACTTCCACTCATCCGCGAGCT 2280  
Db ArgAlaArgSerCysLeuAlaHisSerLeuArgGlnHisPheHisSerIleProProAla 705  
QY 2281 GACCGCGGTGAGCCCAAGAGCGTGCATCCATCCCGGGCTTACTGCTGATATCCGAGAC 2340  
Db ValProGlnGlnLysSerHisAlaMetProGlyPheIleThrLeuIleArgSer 725  
QY 2341 CTGTACAGATGCAAGAGAGCGGCTGCTCGGAGGCTGCAAGCTGATGCTGATGCTG 2400  
Db LeuTyrGlnMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 745  
QY 2401 CAGCTCAAGTGAATTTTCAAGTGTGGCGCGCACTGAGTGTGCTGCTGCTGCTGCTG 2460  
Db HisLeuLysLeuThrPheCysArgValGlyProAlaGlnLysAlaAlaLeuAlaPheVal 765  
QY 2461 CTGAGCACTCCGGCGCGCGCTGCGCGCTGCAAGCTGAGTACAACTGTGGGCTGACAT 2520  
Db LeuGlnHisLeuGlnArgProValAlaLeuGlnLeuLysArgLysHisSerValGlyAspVal 785  
QY 2521 GCGGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580  
Db GlyValGlnGlnLeuArgProCysLeuGlyValCysThrAlaLeuTyrLeuArgAspAsn 805  
QY 2581 AATATCTCAAGCGAGCATCTGCAAGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTG 2640  
Db AsnIleSerPheArgGlyAlaArgThrLeuValGlnCysAlaLeuArgCysGlnGlnLeu 825  
QY 2641 CAGAGTTACTTATTCACAAACAAATGACTGAGCGCTGCAAGCTGCAAGCTGCAAGCTG 2700  
Db GlnLysLeuAlaLeuPheAsnLysLeuThrAspAlaCysAlaCysSerMetAlaLys 845  
QY 2701 CTGCTGATGAGGCGCAAGCTTCTTGGCAATGAGCTGAGGCGCAATGACTCACTGCC 2760  
Db LeuLeuAlaHisLysGlnAsnPheLeuSerLeuArgValIleLysAsnHisIleThrAla 865  
QY 2761 GCGAGACCGCAAGTCTGAGCGAGGCGCTCCAGGCAAGCAAGCTTCTGCAAGTCTG 2820  
Db PheIleArgLysValLeuAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 885  
QY 2821 TTCTGGGCGCAAGAGTGTGAGGCGTGAAGGCGCGCTGCTGCAAGCTTGGTGT 2880  
Db PheIleArgLysValLeuAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 905  
QY 2881 CAGCAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2940  
Db HisGlnAsnLeuLysTyrPheLeuSerValGlnLysAsnHisIleGlySerMetGlyAla 925  
QY 2941 GCTTGGCACTGATGCTGCGCAAGAGCGTACAGAGAACTGCTGCTGAGAGAGAC 3000

```

Db      926 AlaLeuAlaLeuMetLeuGluValAsnLysSerLeuGluGluLeuCysLeuGluGln 945
QY      3001 CATCTCAGATGAGGTGTATGTTCTCTCGAGAGAGACTGAAGAAAAATCAAGTTG 3060
Db      946 HisIleCysAspGluGlyValTyrSerLeuAlaGluGlyLeuLysArgAsnSerThrLeu 965
QY      3061 AAAATCTGAGGTGTCCATTAAGTACCTAGGGGCAAGAGCCCTCTCTGAG 3120
Db      966 LysPheLeuLysLeuSerAspAsnGlyIleThrTyrArgGlyAlaGluAlaLeuGln 985
QY      3121 GCCCCTTAAAGGATGACACCATCTCTGAGAGTCTGGCTCCAGGAGACATTCTCT 3180
Db      986 Ala-LeuSerArgAsnSerAlaIleLeuGluValTyrPheLysArgGlyAsnThrPheSerIle 1005
QY      3181 AAGAGAGGTGACAGCTCGGCTGCAGGAGACACCACTTCTGCT 3226
Db      1005 uGIuGIuIleGlnThrLeuSerSerArgAspAlaArgLeuLeu 1020

```

Search completed: January 7, 2004, 20:05:28  
 Job time : 592.769 secs



and is derived by analysis of the total score distribution.

SUMMARIES

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:04:17 ; Search time 38.0703 seconds  
(without alignments) 4223.507 Million cell updates/sec

Title: US-10-014-269-3

Perfect score: 5304

Sequence: 1 MCSQAFQACRSQQLVLLVS.....NFTSLEEVKLCGRDRLLL 1013

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1930.DAT:\*

2: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2002.DAT:\*

24: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

Result No.	Score	Match	Length	DB	ID	Description
1	5304	100.0	1013	23	AAE25417	Human Nod2 protein
2	5304	100.0	1013	23	ABU04736	Nod2b protein wild
3	5304	100.0	1040	23	AAE25416	Human Nod2 protein
4	5303	100.0	1040	23	ABU04773	Nod2a Y95Y1 proteol
5	5301	99.9	1040	23	ABU04771	Nod2a Y93Y1 proteol
6	5300	99.9	1040	23	ABU04774	Nod2a Y86Y1 proteol
7	5299	99.9	1040	23	AAE25418	Human Nod2 protein
8	5299	99.9	1040	23	ABU04737	Nod2 protein sequ
9	5299	99.9	1040	23	ABU04772	Nod2a N852S proteol
10	5296	99.8	1038	23	ABU04735	Nod2a protein wild
11	5296	99.8	1040	23	ABU04768	Nod2a R702W proteol
12	5296	99.8	1040	23	ABU04769	Nod2a G908R (SNP17
13	5296	99.8	1040	23	ABU04770	Nod2a P268S proteol
14	5296	99.8	1041	22	AAU679119	Amino acid sequenc
15	5291	99.8	1040	23	ABU04776	Nod2a P268S + N852
16	5288	99.7	1040	23	ABU04775	Nod2a P268S + G980
17	5288	99.7	1040	23	ABU04777	Nod2a P268S + R702
18	5128	96.7	1007	23	AAE25434	Human Nod2 mutant.
19	5128	96.7	1007	23	ABU04766	Nod2a protein delt
20	5120	96.5	1007	23	ABU04767	Nod2a P268S + fram
21	4980.5	93.9	1009	23	AAU80865	Human CARD3X prote
22	4174	78.7	795	23	AAU80875	Human CARD3X prote
23	2945	55.5	560	23	AAU80856	Human CARD3X prote
24	2087	39.3	403	23	AAU80870	Human CARD3X ANG10
25	1614	30.4	305	23	AAE25421	Human Nod2 protein
26	1614	30.4	305	23	ABU04740	Nod2 nucleotide bi
27	1564	29.5	296	23	AAU80869	Human CARD3X NB-AR
28	1219.5	23.0	953	20	AAV31145	Murine CARD-4L pro
29	1219.5	23.0	953	22	AAE20004	Mouse CARD-4L (lon
30	1219.5	23.0	953	23	ABG31080	Mouse caspase recr
31	1219.5	23.0	953	23	AAO22136	Protein of murine
32	1219.5	23.0	953	24	ABU56298	Mouse Caspase recr
33	1212	22.9	953	22	AAU79473	Human protein SEQ
34	1209.5	22.8	953	21	AAU15552	Apoptosis related
35	1209.5	22.8	953	22	AAU78489	Human protein SFO
36	1209.5	22.8	953	22	AAE20000	Human CARD-4L (lon
37	1209.5	22.8	953	23	AAO22111	Protein of human C
38	1209.5	22.8	953	24	ABU56273	Human Caspase recr
39	1208.5	22.8	953	23	ABG31076	Human caspase recr
40	1204.5	22.7	953	20	AAV31141	Human CARD-4L prot
41	1171	22.1	966	24	ABU56300	Caspase recruitment
42	1001	18.9	705	21	AAU15559	Apoptosis related
43	1001	18.9	705	21	AAU15562	Fragment of apopto
44	1001	18.9	719	22	AAU75559	Human colon cancer
45	687.5	16.7	779	22	AAU95610	Human protein sequ

Search completed: January 7, 2004, 20:49:51  
Job time : 41.0703 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:54:13 ; Search time 29.7941 Seconds  
(without alignments)  
6851.32 Million cell updates/sec

Title: US-10-014-269-3

Perfect score: 5304

Sequence: 1 MCSQAFQAGRSQVLELVLS.....NPFSLSEYDKLGGCDTLL 1013

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications .AA.\*

1: /cgn2\_6/p/odata/2/pubppa/US07\_PUBCONB.pep.\*  
2: /cgn2\_6/p/odata/2/pubppa/ECI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/p/odata/2/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/p/odata/2/pubppa/US06\_PUBCONB.pep.\*  
5: /cgn2\_6/p/odata/2/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/p/odata/2/pubppa/PCTUS\_PUBCONB.pep.\*  
7: /cgn2\_6/p/odata/2/pubppa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/p/odata/2/pubppa/US08\_PUBCONB.pep.\*  
9: /cgn2\_6/p/odata/2/pubppa/US09A\_PUBCONB.pep.\*  
10: /cgn2\_6/p/odata/2/pubppa/US09B\_PUBCONB.pep.\*  
11: /cgn2\_6/p/odata/2/pubppa/US09C\_PUBCONB.pep.\*  
12: /cgn2\_6/p/odata/2/pubppa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/p/odata/2/pubppa/US10A\_PUBCONB.pep.\*  
14: /cgn2\_6/p/odata/2/pubppa/US10C\_PUBCONB.pep.\*  
15: /cgn2\_6/p/odata/2/pubppa/US10C\_PUBCONB.pep.\*  
16: /cgn2\_6/p/odata/2/pubppa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/p/odata/2/pubppa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/p/odata/2/pubppa/US60\_PUBCONB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	5304	100.0	1013	12	US-10-314-506-3	Sequence 3, Appl1
2	5304	100.0	1013	14	US-10-014-269-3	Sequence 3, Appl1
3	5304	100.0	1013	14	US-10-002-974-3	Sequence 2, Appl1
4	5304	100.0	1040	12	US-10-314-506-2	Sequence 2, Appl1
5	5304	100.0	1040	14	US-10-014-269-2	Sequence 2, Appl1
6	5304	100.0	1040	14	US-10-002-974-2	Sequence 2, Appl1
7	5303	100.0	1040	14	US-10-002-974-67	Sequence 67, Appl1
8	5301	99.9	1040	14	US-10-002-974-63	Sequence 67, Appl1
9	5300	99.9	1040	14	US-10-002-974-69	Sequence 69, Appl1
10	5299	99.9	1040	12	US-10-314-506-4	Sequence 4, Appl1
11	5299	99.9	1040	14	US-10-002-974-65	Sequence 4, Appl1
12	5299	99.9	1040	14	US-10-002-974-65	Sequence 65, Appl1
13	5296	99.8	1040	12	US-10-314-506-57	Sequence 57, Appl1
14	5296	99.8	1040	14	US-10-002-974-57	Sequence 57, Appl1
15	5296	99.8	1040	14	US-10-002-974-59	Sequence 59, Appl1
16	5296	99.8	1040	14	US-10-002-974-61	Sequence 61, Appl1
17	5296	99.8	1041	12	US-10-240-046A-2	Sequence 2, Appl1
18	5291	99.8	1040	14	US-10-002-974-87	Sequence 87, Appl1
19	5288	99.7	1040	12	US-10-314-506-61	Sequence 61, Appl1
20	5288	99.7	1040	14	US-10-002-974-65	Sequence 65, Appl1
21	5288	99.7	1040	14	US-10-002-974-65	Sequence 85, Appl1
22	5288	99.7	1040	14	US-10-002-974-65	Sequence 89, Appl1
23	5128	96.7	1007	12	US-10-314-506-34	Sequence 34, Appl1
24	5128	96.7	1007	14	US-10-014-269-34	Sequence 34, Appl1
25	5128	96.7	1007	14	US-10-002-974-34	Sequence 34, Appl1
26	5120	96.5	1007	12	US-10-314-506-55	Sequence 55, Appl1
27	5120	96.5	1007	14	US-10-002-974-55	Sequence 55, Appl1
28	4980.5	93.9	1009	10	US-09-864-921-107	Sequence 107, App
29	4980.5	93.9	1009	10	US-09-864-921-107	Sequence 188, App
30	4174	79.7	795	10	US-09-864-921-188	Sequence 176, App
31	2945	55.5	560	10	US-09-864-921-188	Sequence 7, Appl1
32	2037	39.3	403	10	US-09-864-921-176	Sequence 7, Appl1
33	1614	30.4	305	12	US-10-014-269-7	Sequence 176, App
34	1614	30.4	305	14	US-10-014-269-7	Sequence 7, Appl1
35	1614	30.4	305	14	US-10-002-974-7	Sequence 7, Appl1
36	1564	29.5	296	10	US-09-864-921-174	Sequence 174, App
37	1219.5	23.0	953	9	US-09-728-721-43	Sequence 43, Appl1
38	1219.5	23.0	953	14	US-10-118-984-43	Sequence 43, Appl1
39	1219.5	23.0	953	15	US-10-295-981-43	Sequence 43, Appl1
40	1209.5	22.8	953	9	US-09-728-721-8	Sequence 8, Appl1
41	1209.5	22.8	953	12	US-10-183-770-4	Sequence 4, Appl1
42	1209.5	22.8	953	14	US-10-105-931-8	Sequence 4, Appl1
43	1209.5	22.8	953	14	US-10-105-931-8	Sequence 8, Appl1
44	1209.5	22.8	953	14	US-10-118-984-8	Sequence 8, Appl1
45	1209.5	22.8	953	15	US-10-013-477-12	Sequence 12, Appl1

Search completed: January 7, 2004, 21:06:00  
Job time : 33.7941 secs

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:49:58 : Search time 16.8633 Seconds

(without alignments)  
5770.121 Million cell updates/sec

Title: US-10-014-269-3

Perfect score: 5304

Sequence: 1 MCSGEAFQAGRSQVLLVLS.....NPFSLSEVDKLGCRDTRLL 1013

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76:\*

2: PIR1:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	408	7.7	1130	2	A48843			MHC class II trans
2	393.5	7.4	1192	2	T17235			hypothetical prote
3	369.5	6.9	1111	2	A59000			matern protein (imp
4	258.5	4.9	456	2	A31857			ribonuclease-inhib
5	237.5	4.5	461	2	A31858			ribonuclease-angio
6	227	4.3	456	2	S20597			ribonuclease-inhib
7	221.5	4.2	535	2	T48102			RAN GTPase activat
8	221	4.2	789	2	T52067			hypothetical prote
9	217	4.1	533	2	T52063			RAN GTPase-activat
10	213	4.0	506	2	A45841			T-complex-associat
11	212.5	4.0	631	2	C89243			protein F28C1.3 (I
12	212.5	4.0	631	2	T21471			hypothetical prote
13	210	4.0	1075	2	T31668			hypothetical prote

14	209.5	3.9	545	2	T52068			RAN GTPase-activat
15	207	3.9	1004	2	T31665			hypothetical prote
16	191.5	3.6	1010	2	T36383			probable large ATP
17	191.5	3.6	1447	2	T42628			neuronal apoptosis
18	184.5	3.5	483	2	S27890			Nasopressin recept
19	173.5	3.3	862	2	T36380			probable large ATP
20	162.5	3.1	332	2	T52069			GTPase-activating
21	160	3.0	1096	2	A55478			neuronal apoptosis
22	157.5	3.0	1119	2	AE2239			proteom disease re
23	156.5	3.0	1096	2	A96607			hypothetical prote
24	155.5	2.9	998	2	T23427			hypothetical prote
25	155	2.9	1411	2	S44370			C2564.3 protein -
26	153.5	2.9	1121	2	T02764			myosin-I binding p
27	150.5	2.8	312	2	B97746			hypothetical prote
28	144.5	2.7	861	2	T36381			probable large ATP
29	144	2.7	662	2	S42799			garp precursor - h
30	143.5	2.7	587	2	UC5300			RAN GTPase activat
31	137	2.6	589	2	T52070			RAN protein homol
32	136	2.6	589	2	A36983			RAN homolog fugu
33	134	2.5	183	2	F97803			hypothetical prote
34	130	2.5	618	2	T48193			hypothetical prote
35	130	2.5	1223	2	E88451			protein K10D2.1 (I
36	129	2.4	1232	2	T05322			hypothetical prote
37	128.5	2.4	386	2	S37691			RAN GTPase activat
38	127.5	2.4	603	2	UC6128			insulin-like growt
39	126.5	2.4	1547	2	A12043			hypothetical prote
40	125.5	2.4	849	2	C97303			hypothetical prote
41	125.5	2.4	1286	2	AB2042			hypothetical prote
42	124.5	2.3	800	2	AB1129			intercalin A (limp
43	124.5	2.3	1088	2	E86312			FL1A6.9 protein -
44	124	2.3	359	1	A42336			tropocollin - hum
45	123.5	2.3	381	2	T43444			hypothetical prote

Search completed: January 7, 2004, 20:50:42  
Job time : 16.2556 secs

GenScore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using SW model

Run on: January 7, 2004, 20:45:37; Search time 33.7667 Seconds  
(without alignments)  
7741.578 Million cell updates/sec

Title: US-10-014-269-3

Perfect score: 5304  
Sequence: 1 MCSQEAFAQSRQVLELVLS.....NTFSLEVDKLGCRPTNLL 1013

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 630525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SPREN1.23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	4185	78.9	1020	11	Q8K320	Q8K320 mus musculus
2	1222.5	23.0	953	11	Q8BH50	Q8BH50 mus musculus
3	1209.5	22.8	953	4	Q81WF5	Q81WF5 homo sapien
4	1019	19.2	706	11	Q8BUT6	Q8BUT6 mus musculus
5	496	9.4	1033	11	Q8R4B8	Q8R4B8 mus musculus
6	479.5	5.0	892	4	Q8NEU4	Q8NEU4 homo sapien
7	425	8.0	1133	7	Q9G3E0	Q9G3E0 rattus norv
8	405.5	7.6	778	4	Q8NF48	Q8NF48 homo sapien
9	401	7.6	1052	7	Q9G3D8	Q9G3D8 rattus norv
10	401	7.6	1073	7	Q9G3D9	Q9G3D9 rattus norv
11	377.5	7.1	713	6	Q9S1Z7	Q9S1Z7 mus musculus
12	377.5	6.8	492	4	Q8TEI3	Q8TEI3 mus musculus
13	362.5	6.8	982	11	Q8BU40	Q8BU40 mus musculus
14	360.5	6.8	733	4	Q8TEE2	Q8TEE2 homo sapien
15	357.5	6.7	1056	4	Q8NF42	Q8NF42 homo sapien
16	344.5	6.5	846	4	Q8IXT0	Q8IXT0 homo sapien
17	332	6.3	692	4	Q96D51	Q96D51 homo sapien
18	321	6.1	660	4	Q8NF06	Q8NF06 homo sapien
19	316.5	6.0	863	11	Q8C6A9	Q8C6A9 mus musculus
20	314	5.9	397	11	Q8BU74	Q8BU74 mus musculus
21	305	5.8	864	7	Q29675	Q29675 homo sapien
22	297	5.6	932	4	Q96K14	Q96K14 homo sapien
23	297	5.6	977	7	Q8HW99	Q8HW99 mus musculus
24	296.5	5.6	312	11	Q8BV65	Q8BV65 mus musculus
25	293	5.5	1097	4	Q9HEV0	Q9HEV0 homo sapien
26	293	5.5	223	4	Q8NI02	Q8NI02 homo sapien
27	284	5.4	605	10	Q93ZV8	Q93ZV8 arabidopsis
28	278	5.2	195	4	Q8NT01	Q8NT01 mus musculus
29	273	5.1	748	11	Q99XMO	Q99XMO mus musculus
30	255	4.8	817	5	Q964V6	Q964V6 dictyostell
31	250.5	4.7	310	11	Q8BYL7	Q8BYL7 mus musculus
32	245.5	4.6	620	16	Q8XZM9	Q8XZM9 ralscopia s
33	243.5	4.6	554	4	Q8H7Z4	Q8H7Z4 homo sapien
34	239.5	4.5	461	4	Q8H7X3	Q8H7X3 homo sapien
35	238.5	4.5	1787	10	Q9V4X9	Q9V4X9 chlamydomon
36	238	4.5	447	4	Q96BD7	Q96BD7 homo sapien
37	237.5	4.5	461	6	Q8H2P9	Q8H2P9 pan troglod
38	237.5	4.5	461	4	Q9S080	Q9S080 homo sapien
39	236.5	4.4	519	11	Q8C249	Q8C249 mus musculus
40	234.5	4.4	825	11	Q8C6M5	Q8C6M5 mus musculus
41	232.5	4.4	753	4	Q969I7	Q969I7 homo sapien
42	232	4.4	681	10	Q9XIK2	Q9XIK2 arabidopsis
43	231.5	4.4	673	11	Q8C0N1	Q8C0N1 mus musculus
44	226.5	4.3	456	11	Q924P4	Q924P4 mus musculus
45	224	4.2				

ALIGNMENTS

RESULT 1  
Q8K320 ID Q8K320 PRELIMINARY; PRT: 1020 AA.  
AC Q8K320;  
DT 01-OCT-2002 (TRENBERG, 22, Created)  
DT 01-OCT-2002 (TRENBERG, 22, Last sequence update)  
DT 01-MAR-2003 (TRENBERG, 23, Last annotation update)

DE CARD15.  
 GN CARD15.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OK NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Tanaga Y., Davey M.P., Martin T.M., Planck S.R., DePriest M.L.,  
 RA Baugh M.M., Suing C., Rosenbaum J.T.;  
 RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15  
 gene";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF520774; AAMT6073.1;  
 DR MGD; MGI:2429397; Card15.  
 DR InterPro; IPR001315; CARD.  
 DR InterPro; IPR007091; LRR\_Nr1nh.  
 DR InterPro; IPR007111; NACHT\_NTPase.  
 DR InterPro; IPR011303; TAD\_DNase.  
 DR PROSITE; PS50209; CARD; 2.  
 DR PROSITE; PS50503; LRR\_Rif; 1.  
 DR PROSITE; PS50837; NACHT; 1.  
 DR PROSITE; PS01337; TAD; 1.  
 SQ SEQUENCE 1020 AA; 113961 MW; 25504905ECF70EBB CRC64;

Query Match 78.9%; Score 4185; DB 11; Length 1020;  
 Best Local Similarity 79.3%; Pred. No. 0;  
 Matches 803; Conservative 82; Mismatches 128; Indels 0; Gaps 0;

QY 1 MCSQAFQARQSDVVELVSSLEGESVYLMLSTVEVWEDYEGFHLLQQLPILHARR 60  
 DB 8 MCSQEEFQARQSLVALLISGLEFSILLWLSVDVLSREDEGLSPQPLSHSARR 67  
 QY 61 LLDVTWNGTWACQKLIAAQEAQADSPKLGQWDPHSIPARLQSHRPAIVRLHS 120  
 DB 68 LLDVTWNGTWACQKLIAAQEAQADSPKLGQWDPHSIPARLQSHRPAIVRLHS 127  
 QY 121 HVENMLDLWEGFVSQYECDEIRLPFTPSQARRLDLATAVYANGLAFAELQHVQELP 180  
 DB 128 HVENMLDLWEGFVSQYECDEIRLPFTPSQARRLDLATAVYANGLAFAELQHVQELP 187  
 QY 181 VPLALPLEAATCKRMKAKRTTVSAQSREFLSTYDGAETLCEDEITYENMLEVWADYAGAG 240  
 DB 188 VPLALPLEAATCKRMKAKRTTVSAQSREFLSTYDGAETLCEDEITYENMLEVWADYAGAG 247  
 QY 241 PRGKSPATLGLELFTPGHINDADTVLVVGEAGSGKSTLLQRLHLLMAAGDPQEEFL 300  
 DB 248 ALQKSPATLGLELFTPGHINDADTVLVVGEAGSGKSTLLQRLHLLMAAGDPQEEFL 307  
 QY 301 VEPFSCQIQCAKAPLSVRLTFEHCQCPDVQGEDIFQLLDHPDNVLLTFDGFDEFKFR 360  
 DB 308 IFFPSCQIQCAKAPLSVRLTFEHCQCPDVQGEDIFQLLDHPDNVLLTFDGFDEFKFR 367  
 QY 361 FTDERHCSPTDPTSVQTLLENLLQGNLKKARVVTSRPAVSALFKRYITCEVILKGF 420  
 DB 368 FTDERHCSPTDPTSVQTLLENLLQGNLKKARVVTSRPAVSALFKRYITCEVILKGF 427

QY 421 SEQGEIYLKRHHPEPVADRLRLIQETSLALGCLHPVSWVWVSKCHQELLQEGSP 480  
 DB 428 SEQGEIYLKRHHPEPVADRLRLIQETSLALGCLHPVSWVWVSKCHQELLQEGSP 487  
 QY 481 KTTTMYLLILQFLLHATPPDSASQGSLLRGRLETLHLGRLALMGKCCYVESA 540  
 DB 488 TSTDMYLLILQFLLHATPPDSASQGSLLRGRLETLHLGRLALMGKCCYVESA 547  
 QY 541 QQLQAQVSPDDISLGFVRAKGVPGSTAPLEFHTFFQCFPAFYALISADVPPALLR 600  
 DB 548 QQLQAQVSPDDISLGFVRAKGVPGSTAPLEFHTFFQCFPAFYALISADVPPALLR 607  
 QY 601 HLENGRGRSPVALLPTVCTQASQGSVALLQALPEPHNLQTTAFLAGLSREHW 660  
 DB 608 HLENGRGRSPVALLPTVCTQASQGSVALLQALPEPHNLQTTAFLAGLSREHW 667  
 QY 661 GLLAECQTSKALLRQACARWCLARSLRHHFHSIPPAVGEAKSVHAPGFTWILRSY 720  
 DB 668 DILAQCVSERVILQROKARACTLAHSIRHHFHSIPPAVGEAKSVHAPGFTWILRSY 727  
 QY 721 EQGBERLARQARGLNVGHKLTFCVGPTECCALAFYQLHRRPVALLQIDVNSVGDIGV 780  
 DB 728 EQGBERLARQARGLNVGHKLTFCVGPTECCALAFYQLHRRPVALLQIDVNSVGDIGV 787  
 QY 781 EQLPCLGVCKALYLRNNNTSDRGICKLEICLHCEQLOKLAFLNNKLTDCGHSKALL 840  
 DB 788 EQLPCLGVCKALYLRNNNTSDRGICKLEICLHCEQLOKLAFLNNKLTDCGHSKALL 847  
 QY 841 AGRQFLALRLANNITPAAGQVLAEGIRGNTSLQFLGFVGNVGECAQALAEALGDHQ 900  
 DB 848 AGRQFLALRLANNITPAAGQVLAEGIRGNTSLQFLGFVGNVGECAQALAEALGDHQ 907  
 QY 901 SIPLSLVGNIGSVGAALATLAAVNLBELICEENHLQDBGVCSLAJGLKNSSLKI 960  
 DB 908 SIPLSLVGNIGSVGAALATLAAVNLBELICEENHLQDBGVCSLAJGLKNSSLKI 967  
 QY 961 LKLSNNGITYLGAELQALENDITLWLNQNTFSLIEVDKLCGRTRILL 1013  
 DB 968 LKLSNNGITYLGAELQALENDITLWLNQNTFSLIEVDKLCGRTRILL 1020

Search completed: January 7, 2004, 20:52:36  
 Job time : 37.7667 secs

and is derived by analysis of the total score distribution.

OM protein - protein search, using sw model

Run on: January 7, 2004, 20:04:17 ; Search time 39.085 seconds  
(without alignments)  
4223.507 Million cell updates/sec

Title: US-10-014-269-2  
Perfect score: 5448  
Sequence: 1 MGEEGASAHDEERASVTL.....NTFSELVKICRDRFLL 1040

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.

Result No.	Score	Query Match	Length	DB ID	Description
1	5448	100.0	1040	23	Human Nod2 protein
2	5447	100.0	1040	23	Nod2a V9351 protel
3	5445	99.9	1040	23	Nod2a V793M protel
4	5444	99.9	1040	23	Nod2a M863V protel
5	5443	99.9	1040	23	Human Nod2 protein
6	5443	99.9	1040	23	Nod2 protein sequ
7	5443	99.9	1040	23	Nod2a N852S protel
8	5440	99.9	1038	23	Nod2a protein wild
9	5440	99.9	1040	23	Nod2a R702W protel
10	5440	99.9	1040	23	Nod2a G908R (SNP)7
11	5440	99.9	1040	23	Nod2a P268S protel
12	5435	99.8	1040	23	Nod2a P268S + N852
13	5432	99.7	1040	23	Nod2a P268S + G980
14	5432	99.7	1040	23	Nod2a P268S + R702
15	5331	97.9	1041	22	Amino acid sequenc
16	5304	97.4	1013	23	Human Nod2 protein
17	5304	97.4	1013	23	Nod2b protein wild
18	5272	96.8	1007	23	Human Nod2 mutant.
19	5272	96.8	1007	23	Nod2a protein delt
20	5264	96.6	1007	23	Nod2a P268S + fram
21	4994.5	91.7	1009	23	Human CARD3X prote
22	4188	76.9	793	23	Human CARD3X prote
23	2959	54.3	560	23	Human CARD3X prote
24	2087	38.3	403	23	Human CARD3X prote
25	1614	29.6	305	23	Human Nod2 protein
26	1614	29.6	305	23	Nod2 nucleotide bi
27	1564	28.7	296	23	Human CARD3X NB-AR
28	1219.5	22.4	953	20	AAV31145
29	1219.5	22.4	953	22	AAV31145
30	1219.5	22.4	953	23	AAV31145
31	1219.5	22.4	953	23	AAV31145
32	1219.5	22.4	953	23	AAV31145
33	1212	22.2	953	22	AAV31145
34	1209.5	22.2	953	21	AAV31145
35	1209.5	22.2	953	22	AAV31145
36	1209.5	22.2	953	22	AAV31145
37	1209.5	22.2	953	23	AAV31145
38	1209.5	22.2	953	24	AAV31145
39	1208.5	22.1	953	23	AAV31145
40	1204.5	22.1	953	20	AAV31145
41	1171	21.5	966	24	AAV31145
42	1001	18.4	705	21	AAV31145
43	1001	18.4	705	21	AAV31145
44	1001	18.4	719	22	AAV31145
45	887.5	16.3	779	22	AAV31145

Search completed: January 7, 2004, 20:49:48  
Job time : 43.085 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: January 7, 2004, 20:54:13 : Search time 30.5882 Seconds  
(without alignments)  
6831.332 Million cell updates/sec

Title: US-10-014-269-2

Perfect score: 5448  
Sequence: 1 MGEEGGASHDEERASVLL.....NTESEVDTGCRDRLLL 1040

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*

1:	/cgn2_6/p/odata/2/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/p/odata/2/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/p/odata/2/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/p/odata/2/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/p/odata/2/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/p/odata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/p/odata/2/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/p/odata/2/pubpaa/US09_PUBCOMB.pep:*
9:	/cgn2_6/p/odata/2/pubpaa/US09_PUBCOMB.pep:*
10:	/cgn2_6/p/odata/2/pubpaa/US09C_PUBCOMB.pep:*
11:	/cgn2_6/p/odata/2/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/p/odata/2/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/p/odata/2/pubpaa/US10A_PUBCOMB.pep:*
14:	/cgn2_6/p/odata/2/pubpaa/US10C_PUBCOMB.pep:*
15:	/cgn2_6/p/odata/2/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/p/odata/2/pubpaa/US10_NEW_PUB.pep:*
17:	/cgn2_6/p/odata/2/pubpaa/US10_NEW_PUB.pep:*
18:	/cgn2_6/p/odata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	DB ID	Description
------------	-------	-------------	-------	-------------

Search completed: January 7, 2004, 21:05:56  
Job time : 33.5882 secs

1	5448	100.0	1040	12	US-10-314-506-2	Sequence 2, Appl
2	5448	100.0	1040	14	US-10-014-269-2	Sequence 2, Appl
3	5448	100.0	1040	14	US-10-002-974-2	Sequence 2, Appl
4	5447	100.0	1040	14	US-10-002-974-67	Sequence 67, Appl
5	5445	99.9	1040	14	US-10-002-974-63	Sequence 63, Appl
6	5444	99.9	1040	14	US-10-002-974-69	Sequence 69, Appl
7	5443	99.9	1040	12	US-10-314-506-4	Sequence 4, Appl
8	5443	99.9	1040	14	US-10-014-269-4	Sequence 4, Appl
9	5443	99.9	1040	14	US-10-002-974-65	Sequence 65, Appl
10	5443	99.9	1040	12	US-10-314-506-57	Sequence 57, Appl
11	5440	99.9	1040	14	US-10-002-974-57	Sequence 57, Appl
12	5440	99.9	1040	14	US-10-002-974-59	Sequence 59, Appl
13	5440	99.9	1040	14	US-10-002-974-61	Sequence 61, Appl
14	5440	99.9	1040	14	US-10-002-974-63	Sequence 63, Appl
15	5435	99.8	1040	14	US-10-002-974-67	Sequence 67, Appl
16	5432	99.7	1040	12	US-10-314-506-59	Sequence 59, Appl
17	5432	99.7	1040	12	US-10-314-506-61	Sequence 61, Appl
18	5432	99.7	1040	14	US-10-002-974-85	Sequence 85, Appl
19	5432	99.7	1040	14	US-10-002-974-89	Sequence 89, Appl
20	5331	97.9	1041	12	US-10-240-046A-2	Sequence 2, Appl
21	5304	97.4	1013	12	US-10-314-506-3	Sequence 3, Appl
22	5304	97.4	1013	14	US-10-014-269-3	Sequence 3, Appl
23	5304	97.4	1013	14	US-10-002-974-3	Sequence 3, Appl
24	5272	96.8	1007	12	US-10-314-506-34	Sequence 34, Appl
25	5272	96.8	1007	14	US-10-014-269-34	Sequence 34, Appl
26	5272	96.8	1007	14	US-10-002-974-34	Sequence 34, Appl
27	5264	96.6	1007	12	US-10-314-506-55	Sequence 55, Appl
28	5264	96.6	1007	14	US-10-002-974-55	Sequence 55, Appl
29	4994.3	91.7	1009	10	US-09-864-921-107	Sequence 107, App
30	4188	76.9	795	10	US-09-864-921-188	Sequence 188, App
31	2959	54.3	560	10	US-09-864-921-83	Sequence 83, Appl
32	2087	38.3	403	10	US-09-864-921-176	Sequence 176, App
33	1614	29.6	305	12	US-10-314-269-7	Sequence 7, Appl
34	1614	29.6	305	14	US-10-014-269-7	Sequence 7, Appl
35	1614	29.6	305	14	US-10-002-974-7	Sequence 7, Appl
36	1564	28.7	296	10	US-09-864-921-174	Sequence 174, App
37	1219.5	22.4	953	9	US-09-728-721-43	Sequence 43, Appl
38	1219.5	22.4	953	14	US-10-118-984-43	Sequence 43, Appl
39	1219.5	22.4	953	15	US-10-295-981-43	Sequence 43, Appl
40	1209.5	22.2	953	9	US-09-728-721-8	Sequence 8, Appl
41	1209.5	22.2	953	12	US-10-028-374-4	Sequence 4, Appl
42	1209.5	22.2	953	12	US-10-183-770-4	Sequence 4, Appl
43	1209.5	22.2	953	14	US-10-105-931-8	Sequence 6, Appl
44	1209.5	22.2	953	14	US-10-118-984-8	Sequence 8, Appl
45	1209.5	22.2	953	15	US-10-013-477-12	Sequence 12, Appl

OK protein - protein search, using sw model

Run on: January 7, 2004, 20:49:58 ; Search time 17.333 Seconds

(Without alignments)  
3770.121 Million cell updates/sec

Title: US-10-014-269-2

Perfect score: 5448

Sequence: 1 MGEEGSSASIDEERASVLT.....NTPSLEVDKLGCRDTRILL 1040

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	408	7.5	1130	2	A48843		MHC class II trans
2	393.5	7.2	1192	2	T11725		hypothetical prote
3	369.5	6.7	1111	2	A59000		matern protein [imp
4	258.5	4.7	456	2	A31657		ribonuclease inhib
5	237.5	4.4	461	2	A31658		ribonuclease-angio
6	227	4.2	456	2	S20597		ribonuclease inhib
7	221.5	4.1	535	2	T48102		RAN GTPase activat
8	221	4.1	789	2	T52067		hypothetical prote
9	217	4.0	553	2	T52063		ran GTPase-associat
10	213	3.9	506	2	A45841		T-complex-associat
11	212.5	3.9	631	2	C89243		protein F28C1.3 [i
12	212.5	3.9	631	2	T21471		hypothetical prote
13	210	3.9	1075	2	T31668		hypothetical prote

14	209.5	3.8	545	2	T52068		RAN GTPase-activat
15	207	3.8	1004	2	T31665		hypothetical prote
16	196	3.6	1447	2	T42628		neuronal apoptosis
17	191.5	3.5	1010	2	T36383		probable large ATP
18	184.5	3.4	483	2	S27880		Nasopressin recept
19	173.5	3.2	862	2	T36380		probable large ATP
20	162.5	3.0	332	2	T52069		GTPase-activating
21	160	2.9	1232	2	A55478		neuronal apoptosis
22	157.5	2.9	1096	2	A96607		protein disease re
23	156.5	2.9	1119	2	AB2239		hypothetical prote
24	155.5	2.9	998	2	T23427		C2964.3 protein -
25	155	2.8	1411	2	S44770		hypothetical prote
26	153.5	2.8	1121	2	T02764		myosin-I binding p
27	150.5	2.8	312	2	B97746		hypothetical prote
28	144.5	2.7	861	2	T36381		probable large ATP
29	144	2.6	662	2	S42799		gasp precursor - h
30	143.5	2.6	587	2	UC5300		RAN GTPase activat
31	137	2.5	589	2	T52070		RNAPI protein homol
32	136	2.5	589	2	A36883		RNAPI homolog fud1
33	134	2.5	183	2	F97803		hypothetical prote
34	130	2.4	418	2	T48193		hypothetical prote
35	130	2.4	1223	2	E88451		protein K10D2.1 [i
36	129	2.4	1232	2	T05322		hypothetical prote
37	128.5	2.4	386	2	S37691		insulin-like growt
38	127.5	2.3	603	2	UC6128		hypothetical prote
39	126.5	2.3	1547	2	A12043		hypothetical prote
40	125.5	2.3	849	2	C97303		hypothetical prote
41	125.5	2.3	1256	2	AB2042		hypothetical prote
42	124.5	2.3	800	2	AB1129		Internalin A [limp
43	124.5	2.3	1088	1	E86312		FluA6.9 protein -
44	124	2.3	359	1	A42336		tropomodulin - hum
45	124	2.3	1151	2	T30936		reverse transcript

Search completed: January 7, 2004, 20:50:37  
Job time : 16.5556 secs



OM protein - protein search, using sw model

Run on: January 7, 2004, 20:45:37 ; Search time 34.6667 Seconds  
(without alignments)  
7741.578 Million cell updates/sec

Title: US-10-014-269-2

Perfect score: 5448  
Sequence: 1 MGEEGGASHDEERAVYL.....NTFSLEVDKGRDRLLL 1040

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries  
SPREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_podent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	4195	77.0	1020	11	Q8K320	Q8K320 mus musculus
2	1222.5	22.4	953	11	Q8B80	Q8B80 mus musculus
3	1209.5	22.2	953	11	Q8B80	Q8B80 mus musculus
4	1019	18.7	706	11	Q8B06	Q8B06 mus musculus
5	496	9.1	1033	11	Q8B48	Q8B48 mus musculus
6	479.5	8.8	892	4	Q8NEU4	Q8NEU4 mus musculus
7	425	7.8	1153	7	Q8GJEO	Q8GJEO mus musculus
8	425	7.8	1155	7	Q8PPI1	Q8PPI1 mus musculus
9	405.5	7.4	1778	4	Q8NFD8	Q8NFD8 mus musculus
10	401	7.4	1052	7	Q8GJDE	Q8GJDE mus musculus
11	401	7.4	1073	7	Q8GJDN	Q8GJDN mus musculus
12	377.5	6.9	713	6	Q8S1Z7	Q8S1Z7 mus musculus
13	362.5	6.7	499	4	Q8RT13	Q8RT13 mus musculus
14	360.5	6.6	982	11	Q8BU00	Q8BU00 mus musculus
15	357.5	6.6	733	4	Q8TEE2	Q8TEE2 mus musculus
16	344.5	6.3	1056	4	Q8N442	Q8N442 mus musculus
17	332	6.1	846	4	Q8IXT0	Q8IXT0 mus musculus
18	321	5.9	692	4	Q8NFD5	Q8NFD5 mus musculus
19	316.5	5.8	660	4	Q8NFD6	Q8NFD6 mus musculus
20	314	5.8	863	11	Q8C639	Q8C639 mus musculus
21	305	5.6	397	11	Q8B0F4	Q8B0F4 mus musculus
22	297	5.5	884	7	Q29675	Q29675 mus musculus
23	297	5.5	932	4	Q86K14	Q86K14 mus musculus
24	296.5	5.4	977	7	Q8H959	Q8H959 mus musculus
25	293	5.4	312	11	Q8BV65	Q8BV65 mus musculus
26	293	5.4	1097	4	Q8H6Y0	Q8H6Y0 mus musculus
27	284	5.2	223	4	Q8N102	Q8N102 mus musculus
28	279	5.1	605	10	Q93ZV8	Q93ZV8 mus musculus
29	273	5.0	195	4	Q8N101	Q8N101 mus musculus
30	255	4.7	748	11	Q93H60	Q93H60 mus musculus
31	250.5	4.6	817	5	Q964V6	Q964V6 mus musculus
32	245.5	4.5	310	11	Q8BY17	Q8BY17 mus musculus
33	243.5	4.5	620	16	Q8XZ99	Q8XZ99 mus musculus
34	239.5	4.4	554	4	Q8H724	Q8H724 mus musculus
35	238.5	4.4	461	4	Q8ITX8	Q8ITX8 mus musculus
36	238	4.4	1787	10	Q8K4X9	Q8K4X9 mus musculus
37	237.5	4.4	447	4	Q86ED7	Q86ED7 mus musculus
38	237.5	4.4	461	6	Q8H2P9	Q8H2P9 mus musculus
39	236.5	4.3	461	4	Q8B080	Q8B080 mus musculus
40	234.5	4.3	519	11	Q8C249	Q8C249 mus musculus
41	232.5	4.3	825	11	Q8C6M5	Q8C6M5 mus musculus
42	232	4.3	753	4	Q869L7	Q869L7 mus musculus
43	231.5	4.2	681	10	Q8Y1K2	Q8Y1K2 mus musculus
44	229	4.2	1403	11	Q8CH64	Q8CH64 mus musculus
45	228	4.2	1403	11	Q8CH68	Q8CH68 mus musculus

ALIGNMENTS

RESULT 1  
Q8K320 ID Q8K320 PRELIMINARY: PRT: 1020 AA.  
AC Q8K320,  
DT 01-OCT-2002 (TRENBLER: 22, Created)  
DT 01-OCT-2002 (TRENBLER: 22, Last sequence update)  
DT 01-MAR-2003 (TRENBLER: 23, Last annotation update).

DE CARD15.  
 GN Mus musculus (House).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c.  
 RA Tanaga Y., Davey M.P., Martin T.M., Planck S.R., Defriest M.L.,  
 RA Baugh M.M., Suling C., Rosenbaum J.T.;  
 RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15  
 RT gene";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF520774; AM076073.1; -  
 DR MGI; MGI:2429397; Card15.  
 DR InterPro; IPR001315; CARD.  
 DR InterPro; IPR007091; LRR\_RNinh.  
 DR InterPro; IPR007111; NACHT\_NTPase.  
 DR InterPro; IPR001301; TARD\_DNase.  
 DR PROSITE; PS50209; CARD\_2.  
 DR PROSITE; PS50503; LRR\_Rif\_1.  
 DR PROSITE; PS50837; NACHT\_1.  
 DR PROSITE; PS01137; TARD\_1.  
 DR PROSITE; PS01137; TARD\_1.  
 SQ SEQUENCE 1020 AA; 11361 MM; 25504905 ECF70FEB CRC64;

Query Match 77.0%; Score 4196; DB 11; Length 1020;  
 Best Local Similarity 79.2%; Pred. No. 0;  
 Matches 804; Conservative 83; Mismatches 128; Indels 0; Gaps 0;

QY 26 CEMCQEAFOAQSSQVLELVSSGSLGFEESVLDKLSWEVLSWEDVEGHLLQPLSHLA 85  
 DB 6 CDMCQEEFEQARSQVLAALLSSLEGESLDMVLSVDSREDEGLSPGQPLSHSA 65  
 QY 86 RLIDTWNKGTWACQKTLAAQEAQADSPKLGKGMDSHPHLPARDIOSHPATVRL 145  
 DB 66 RLIDTWNKGTWACQKTLAAQEAQADSPKLGKGMDSHPHLPARDIOSHPATVRL 125  
 QY 146 HSHVENMLDLAMERGVSCYECDEIRLPIFTPSQARRLLDLATVANGLAFLQWQE 205  
 DB 126 YNHEVEMLELAREGFLSYECEIRLPIFTSSQARRLLDLAAVANGLAFLQWVE 185  
 QY 206 LPVPLALPLEATCKRYAKLRTTVAQSRFTSGAETICLEDIYENTLEVMADVM 265  
 DB 186 LPAPLPLEYBAECQKFTSKRTMTLQSRFTSTYDSSNLCLEDIYENTLELDQEVGT 245  
 QY 266 AGPQKSPATLGLLEIFSTPHLNDADTVLVVGEAGSGKSTLLQRLHLMAAGQDFEF 325  
 DB 246 AGALQKSPAILGLEDFSTPHLNDADTVLVVGEAGSGKSTLLQRLHLMAAGSGDFEF 305  
 QY 326 LFPVFPSCROLOCAKPLSVRTILFEHCMPDVQGEDIFQLLDHPDRVLTDFGDEFK 385  
 DB 306 LFPVFPSCROLOCAKPLSVRTILFEHCMPDVQGEDIFQLLDHPDRVLTDFGDEFK 365  
 QY 386 FRTDREHCSPTDPTSVOTILFNLLQNLKLNARKVVTSPAAVSAPFLKATRTENLK 445  
 DB 366 FRTDREHCSPTDPTSVOTILFNLLQNLKLNARKVVTSPAAVSAPFLKATRTENLK 425

QY 446 GFSEQIEIYLRKREHEPGYADRLRLIQTSLAHGCHLPVFSWVW/SKCHQELLQDESG 505  
 DB 426 GFSEQIEIYLRKREHEPGYADRLRLIQTSLAHGCHLPVFSWVW/SKCHQELLQDESG 485  
 QY 506 SPKTTDMYLLIQEFLKATPPDASAGGLPGLRGLPFLHLGLALWGLMKCYVF 565  
 DB 486 SPKTTDMYLLIQEFLKATPPDASAGGLPGLRGLPFLHLGLALWGLMKCYVF 545  
 QY 566 SAQQLAAQVSPDDISLGFVTRAKGVGSTAPLEFLHTPOCFPAAPYLAASDVAPL 625  
 DB 546 SAQQLAAQVSPDDISLGFVTRAKGVGSTAPLEFLHTPOCFPAAPYLAASDVAPL 605  
 QY 626 LKHLNCGRPQNSPPARILPTMCIOASGKSSVALLQXAEPRNLQITAAFLAGLSRE 685  
 DB 606 LKHLNCGRPQNSPPARILPTMCIOASGKSSVALLQXAEPRNLQITAAFLAGLSRE 665  
 QY 686 HMGLLAECOTSEKALLRHOACAPWCIASTRKHFTSIPPAAPGEAKSVHAMPFTWILRS 745  
 DB 666 HMGLLAECOTSEKALLRHOACAPWCIASTRKHFTSIPPAAPGEAKSVHAMPFTWILRS 725  
 QY 746 LYBQEEERLAKAAGLVNGHLKTFCSVSPTECAALAFVQLHRRPVYALQLDVNSVGI 805  
 DB 726 LYBQEEERLAKAAGLVNGHLKTFCSVSPTECAALAFVQLHRRPVYALQLDVNSVGI 785  
 QY 806 GVEQLLPGLGVCKALYLRDNNTSDRGICKEICQLHCEQLOKIALPNKKLIDGCAHSAK 865  
 DB 786 GVEQLLPGLGVCKALYLRDNNTSDRGICKEICQLHCEQLOKIALPNKKLIDGCAHSAK 845  
 QY 866 LACRQNFALRLANNVITTAAGQVLAEGLRGNTSLQPLGFQGNRVGDEGAQALAEALGD 925  
 DB 846 LACRQNFALRLANNVITTAAGQVLAEGLRGNTSLQPLGFQGNRVGDEGAQALAEALGD 905  
 QY 926 HOSLWSTLVGNNTSGVGAQALALALANVMLEELCBENHLQDEGVCSLAEGKRNSTL 985  
 DB 906 HOSLWSTLVGNNTSGVGAQALALALANVMLEELCBENHLQDEGVCSLAEGKRNSTL 965  
 QY 986 KTKLSNNCITYLGAELALQALERNDITLFWLRGNTFSLSEEDVXKLGCRDTRILL 1040  
 DB 966 KTKLSNNCITYLGAELALQALERNDITLFWLRGNTFSLSEEDVXKLGCRDTRILL 1020

Search completed: January 7, 2004, 20:52:32  
 Job time : 39.6667 secs

OK nucleic - protein search, using frame\_plus\_nzp model

Run on: January 7, 2004, 19:23:16 ; Search time 110.738 Seconds

(without alignments)  
12857.211 Million cell1 updates/sec

Title: US-10-014-269-1

Perfect score: 8270

Sequence: 1 gtagacagatccagctcac.....ataacgttgatcaaac 4485

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 150 summaries

Command line parameters:

-MODEL=frame+npz.model -DEV=xlp

Q=/gen2\_1/USPRO\_spool\_p/US10014269/runat\_07012004\_175228\_24497/app\_query.fasta\_1

.9358

-DB=A\_Geneseq\_19Jun03 -OPMT=fastan -SUFFIX=reg -MINMATCH=0.1 -LOOPCL=0

-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi

-LIST=150 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50

-MODE=LOCAL -OUTFMT=pro -NORTEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10014269 -GCN 1.1 311 -runat\_07012004\_175228\_24497 -NCPUs=6 -ICPU=3

-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGID=0

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database:

A\_Geneseq\_19Jun03:  
1: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1980.DAT:  
2: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT:  
3: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT:  
4: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT:  
5: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT:  
6: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT:  
7: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT:  
8: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT:  
9: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT:  
10: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5448	65.9	1040	23	AAE25416 Human Nod2 protein
2	5447	65.9	1040	23	ABJ04773 Nod2a V9551 protei
3	5445	65.8	1040	23	ABJ04771 Nod2a W679M protei
4	5444	65.8	1040	23	ABJ04774 Nod2a N865V protei
5	5443	65.8	1040	23	AAE25418 Human Nod2 protein
6	5443	65.8	1040	23	ABJ04737 Nod2a protein sequ
7	5443	65.8	1040	23	ABJ04772 Nod2a N852S protei
8	5440	65.8	1038	23	ABJ04735 Nod2a R702W protei
9	5440	65.8	1040	23	ABJ04768 Nod2a P268S (SNP17
10	5440	65.8	1040	23	ABJ04769 Nod2a P268S + N852
11	5440	65.8	1040	23	ABJ04770 Nod2a P268S + N852
12	5432	65.7	1040	23	ABJ04775 Nod2a P268S + R702
13	5432	65.7	1040	23	ABJ04776 Nod2a P268S + R702
14	5432	65.7	1040	23	AAE29419 Amino acid sequenc
15	5331	64.5	1041	22	AAE29417 Human Nod2 protein
16	5304	64.1	1013	23	ABJ04736 Nod2b protein wild
17	5304	64.1	1013	23	AAE25434 Human Nod2 mutant
18	5272	63.7	1007	23	ABJ04766 Nod2a P268S + fram
19	5272	63.7	1007	23	ABJ04767 Nod2a P268S + fram
20	5264	63.7	1007	23	ABJ04766 Nod2a P268S + fram
21	4994.5	60.4	1009	23	AAU80875 Human CARD3X prote
22	4188	50.6	795	23	AAU80875 Human CARD3X prote
23	2959	35.8	560	23	AAU80856 Human CARD3X prote
24	2087	25.2	403	23	AAE25421 Human Nod2 protein
25	1614	19.5	305	23	AAE25421 Human Nod2 protein
26	1614	19.5	305	23	AAE25421 Human Nod2 protein
27	1564	18.9	296	23	AAU80869 Murine CARD-4L pro
28	1219.5	14.7	953	22	ABG31084 Mouse Caspase recr
29	1219.5	14.7	953	22	ABG31084 Mouse Caspase recr
30	1219.5	14.7	953	22	AAO22136 Protein of murine
31	1219.5	14.7	953	22	ABU56288 Mouse Caspase recr
32	1219.5	14.7	953	22	ABU56288 Mouse Caspase recr
33	1212	14.7	959	22	AAU79473 Human protein SEQ

34	1209.5	14.6	953	21	AAH15552	Apoptosis related	91	433	5.2	1130	19	AAW80313	CITF (class II tr
35	1209.5	14.6	953	22	AAH78489	Human protein SQ	92	433	5.2	1130	19	AAW57056	Class II trans act
36	1209.5	14.6	953	23	AAH20080	Human CARD-4L (lon	93	433	5.2	1207	19	AAW80311	CITF (class II tr
37	1209.5	14.6	953	24	AAH22111	Protein of human C	94	429.5	5.2	896	23	AAH05592	Human PYRIN-11 pro
38	1209.5	14.6	953	25	AAH62823	Human Caspase recr	95	425	5.1	320	23	AAH04763	Nod1 protein seque
39	1208.5	14.6	953	26	ABG31076	Human Caspase recr	96	421.5	5.1	869	22	AAH63318	Human transcriptio
40	1204.5	14.6	953	27	AAH31141	Human CARD-4L prot	97	418	5.1	81	23	AAH08068	Human CARD3X CARDB
41	1171	14.2	966	28	AAH56300	Caspase recruitment	98	415	5.0	1454	22	AAH72673	Human NAC beta iso
42	1001	12.1	705	29	AAH15559	Apoptosis related	99	414.5	5.0	1429	23	AAH7855	Human caspase recr
43	1001	12.1	705	30	AAH15562	Fragment of apopto	100	414.5	5.0	1429	23	AAH17855	Human CARD-4L
44	1001	12.1	719	31	AAH55889	Human colon cancer	101	414.5	5.0	1473	22	AAH67558	Human G-protein co
45	887.5	10.7	779	32	AAH55890	Human protein sequ	102	411.5	5.0	1473	22	AAH72674	Human CARD-4L
46	858	10.4	166	33	ABG49110	Human liver peptid	103	411	5.0	830	19	AAH50314	Human NAC gamma or
47	858	10.4	166	34	ABG49110	Human liver peptid	104	411	5.0	1397	22	AAH72671	CITF (class II tr
48	858	10.4	166	35	ABH19702	Protein #1701 enco	105	411	5.0	1442	22	AAH72671	Human NB-ARC and C
49	858	10.4	166	36	AAH55062	Human brain expres	106	410.5	5.0	1429	23	AAH62571	Human CARD-7 poly
50	858	10.4	166	37	AAH67453	Human bone marrow	107	410.5	5.0	1429	23	AAH62571	Human leucine rich
51	858	10.4	166	38	AAH30324	Human bone marrow	108	410.5	5.0	1429	23	AAH62571	Human leucine rich
52	858	10.4	166	39	ABG37074	Human liver peptid	109	410.5	5.0	1429	23	AAH62571	Human leucine rich
53	840	10.2	155	40	ABG58852	Human liver peptid	110	410.5	5.0	1429	23	AAH62571	Human leucine rich
54	840	10.2	155	41	ABH43475	Protein #10981 enco	111	410.5	5.0	1429	23	AAH62571	Human leucine rich
55	840	10.2	155	42	ABH66439	Human liver peptid	112	410.5	5.0	1473	22	AAH72673	Human caspase recr
56	840	10.2	155	43	AAH64407	Human brain expres	113	410.5	5.0	1473	22	AAH72673	Human caspase recr
57	840	10.2	155	44	AAH72225	Human bone marrow	114	408	4.9	76	23	AAH72711	Human NAC beta iso
58	840	10.2	155	45	AAH21157	Protein #1591 enco	115	399.5	4.8	919	23	AAH43483	Human CARD3X CARDA
59	840	10.2	155	46	AAH37368	Protein #11405 enco	116	399.5	4.8	919	23	AAH43483	Human CARD3X CARDA
60	840	10.2	155	47	ABG46236	Human liver peptid	117	396.5	4.8	920	23	AAH53258	NOVX related prote
61	820	9.9	146	48	ABG58853	Human liver peptid	118	396	4.8	858	23	AAH53258	Human NODD-13 prot
62	820	9.9	146	49	ABH43476	Protein #10982 enco	119	396	4.8	858	23	AAH53258	Human nucleotide b
63	820	9.9	146	50	ABH66440	Human brain expres	120	396	4.8	994	23	AAH15593	Human PYRIN-3 prot
64	820	9.9	146	51	AAH64408	Human bone marrow	121	395	4.8	952	22	ABG03924	Novel human diapo
65	820	9.9	146	52	AAH72226	Protein #11406 enco	122	392	4.8	19338	24	AAH17862	Pyrim domain conta
66	820	9.9	146	53	AAH21158	Human brain expres	123	391.5	4.7	1022	23	AAH17862	Pyrim domain conta
67	820	9.9	146	54	AAH37368	Protein #11406 enco	124	390.5	4.7	1049	22	AAH65894	Human CGD seq ID
68	820	9.9	146	55	ABG46237	Human bone marrow	125	385	4.7	980	23	AAH65894	Human nucleotide b
69	652.5	7.9	180	56	AAH80876	Human CARD3X #1.	126	384.5	4.6	1851	22	AAH00023	Human activated T-
70	593	7.2	140	57	AAH80877	Human CARD3X #2.	127	384.5	4.6	1851	22	AAH00023	Human activated T-
71	588	7.1	139	58	AAH80878	Human CARD3X #3.	128	384.5	4.6	1851	22	AAH00023	Human activated T-
72	531	6.4	1061	59	AAH15590	Human CARD-4L part	129	379.5	4.6	1062	22	AAH65895	Pyrim domain conta
73	530.5	6.4	490	60	AAH31142	Human CARD-4L part	130	379.5	4.6	1062	22	AAH65895	Pyrim domain conta
74	530.5	6.4	490	61	ABG1077	Human CARD-4L (sho	131	375	4.5	1033	23	AAH7856	Human nucleotide b
75	530.5	6.4	490	62	ABG1077	Human CARD-4L (sho	132	375	4.5	1033	23	AAH7856	Human nucleotide b
76	530.5	6.4	490	63	AAH22127	Human caspase recr	133	375	4.5	1033	23	AAH7856	Human nucleotide b
77	530.5	6.4	490	64	ABH56289	Protein of human C	134	368	4.5	1033	24	AAH7856	Human nucleotide b
78	530.5	6.4	1089	65	AAH17857	Human CARD-4L	135	366.5	4.5	19338	24	AAH7856	Human nucleotide b
79	522	6.3	97	66	AAH23419	Human CARD-4L	136	366	4.4	740	22	AAH7856	Human nucleotide b
80	522	6.3	97	67	AAH23419	Human CARD-4L	137	365.5	4.4	1111	23	AAH7856	Human nucleotide b
81	497	6.0	1034	68	AAH08503	Human CARD-4L	138	364	4.4	1344	23	AAH15585	Human CARD-4L
82	497	6.0	1034	69	AAH08503	Human CARD-4L	139	363	4.4	19338	24	AAH7856	Human CARD-4L
83	485	5.9	94	70	AAH25420	Human CARD-4L	140	360.5	4.4	681	24	AAH7856	Human CARD-4L
84	485	5.9	94	71	AAH25420	Human CARD-4L	141	358.5	4.3	2312	23	AAH7856	Human CARD-4L
85	442.5	5.4	975	72	AAH72596	Human CARD-4L	142	357	4.3	1033	23	AAH7856	Human CARD-4L
86	436.5	5.3	19338	73	AAH72596	Human CARD-4L	143	357	4.3	1162	23	AAH7856	Human CARD-4L
87	433	5.2	987	74	AAH72596	Human CARD-4L	144	355.5	4.3	1200	23	AAH7856	Human CARD-4L
88	433	5.2	1106	75	AAH72596	Human CARD-4L	145	355.5	4.3	1200	24	AAH7856	Human CARD-4L
89	433	5.2	1130	76	AAH72596	Human CARD-4L	146	355	4.3	19338	24	AAH7856	Human CARD-4L
90	433	5.2	1130	77	AAH72596	Human CARD-4L	147	352.5	4.3	19338	24	AAH7856	Human CARD-4L

148 350.5 4.2 1111 24 AAE31748 Mouse WATER protel  
c 149 348 4.2 19938 24 AAB86398 Streptomyces virid  
c 150 343.5 4.2 777 14 AAR37740 Collagen-like poly

ALIGNMENTS

RESULT 1

ID AAE25416 standard; Protein; 1040 AA.

XX AAE25416;

DT 30-OCT-2002 (first entry)

DE Human Nod2 protein #1.

KW Human; Nod2 polypeptide; gene shuffling; transgenic plant; apoptosis;  
KW Crohn's disease; antiinflammatory; gene therapy.

XX Homo sapiens.

PV W0200236773-A2.

PD 10-MAY-2002.

PF 26-OCT-2001; 2001WO-US47699.

PR 30-OCT-2000; 2000US-244289P.

PR 26-OCT-2001; 2001US-0244289.

XX (UNMT ) UNIV MICHIGAN.

PI Nunez G, Itohara N, Ogura Y;

DR WPI; 2002-557300/57.

DR N-PSDB; AAD41434.

PT Novel Nod2 polypeptides and polynucleotides useful in treating diseases

PT or altering physiological states marked by apoptosis of cells or other

PT nuclear factor-kappa B mediated effects -

XX Claim 25; Fig 13; 18pp; English.

CC The invention relates to a purified Nod2 polypeptide and its  
CC polynucleotide. The invention is useful for inhibiting Nod2 interaction  
CC with NF-kB and its associated proteins. Nod2 polynucleotides used in  
CC frame to a marker sequence is useful for purifying Nod2 polypeptide.  
CC Nod2 polynucleotides are also useful in gene shuffling or sexual  
CC polymerase chain reaction (PCR) procedures. Nod2 polynucleotides are  
CC also useful to either increase or decrease the level of Nod2 mRNA and/or  
CC protein in transfected cells as compared to the levels in wild-type cells  
CC and to decrease the level of Nod2 mRNA or protein in transgenic plants,  
CC plant tissues and plant cells as compared to the levels in wild-type.  
CC The invention is useful for inhibiting NF-kB activity and thus for  
CC treating Crohn's disease. Nod2 polypeptides and polynucleotides are  
CC useful in treating diseases or altering physiological states marked by

CC apoptosis of cells or other NF-kB mediated effects. Nod2 gene is useful  
CC in gene therapy. The present sequence is Nod2 protein.  
XX  
SQ Sequence 1040 AA;

Alignment Scores:

Pred. No.: 0 Length: 1040  
Score: 3448.00 Matches: 1040  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 65.88% Indels: 0  
DB: 23 Gaps: 0

US-10-014-269-1 (1-4485) x AAE25416 (1-1040)

QY 106 ATGGGGAGAGAGGGGTGTTCAAGCTCTCAAGATGAGAGAAAGCAAGTCTCTCTC 165  
DB 1 MetGlyGluGluGlyGlySerAlaSerHisAlaSpGluGluGlyAlaSerValLeuLeu 20  
QY 166 GAGCATCTCCGGGTGTGAATGCTCCAGAGAGGCTTTCAAGCAAGAGAGCCAG 225  
DB 21 GlyHisSerProGlyGlyGluValCysSerGlnGluAlaGlnAlaGlySerGln 40  
QY 226 CTGATCAAGCTGCTGGTCTCAAGAGTCCCTGGAAGCTTCAAGAGTCTGAGTGGCTG 285  
DB 41 LeuValGluLeuLeuValSerGlySerLeuGluGlyPheGluSerValLeuLeuPheLeu 60  
QY 286 CTGCTCTGGAGGCTCTCTCTGGAGAGATCAAGAGGCTTCAAGTCTGGAGGCTGCT 345  
DB 61 LeuSerTrpValLeuSerTrpGluGlyGlyPheHisLeuLeuGlyGlnPro 80  
QY 346 CTCTCCCACTGGCCAGGCGCTCTGGACACCGTCTGAATAAAGGTACTGGGCTGT 405  
DB 81 LeuSerHisLeuAlaAlaGlyAlaGluLeuLeuPheTrpValTrpAsnGlyTrpAlaCys 100  
QY 406 CAGAAAGCTCATGCGCGCTGCCCAAGAGAGCCAGAGCCAGTCCCAAGCTGCAT 465  
DB 101 GlnGlyLeuValAlaAlaAlaGlnGluAlaGlnAlaAlaSerGlnSerPheGlyLeuHis 120  
QY 466 GAGTGGTGGAGCCCGCACTGGCTCCAGCCAGCCGAGTCAAGTCAAGCCGAGCC 525  
DB 121 GlyCysTrpAspPheHisSerLeuHisPheAlaGlyAspLeuGlnSerHisArgProAla 140  
QY 526 ATTGTCAAGAGCTCCAGAGCCATGTGAGAAATGCTGGACCTGGCATGGAGCGGGGT 585  
DB 141 IleValArgArgLeuHisSerHisValGluAlaSerHisLeuAlaTrpGluArgGly 160  
QY 586 TTCCTCAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 645  
DB 161 PheValSerPheTrpGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 180  
QY 646 GCAAGAGAGGCTGCTGATCTGCCACGGTGAAGAGCAAGTGGCTGGCTTCTCTCTA 705  
DB 181 AlaArgTrpGluLeuAlaLeuAlaLeuValValAlaAsnGlyLeuAlaAlaPheLeuLeu 200  
QY 706 CAAGATGTCAAGAAATTACAGTCCCATGGCCCTGCTTGGAGTGCATGCAAG 765

Db 201 GlnHisValGlnGluLeuProValProLeuAlaLeuProLeuGluAlaAlaThrCysGlys 220  
QY 766 AAGTATATGGCCAAAGCTGAGACCAAGGTGTCTGCTCAAGTCTGCTCTCAAGTACTAT 825  
Db 221 LysTyrMetCalaLysLeuArgThrThrValSerAlaGlnSerArgPheLeuSerThrTyr 240  
QY 826 GATGAGACAGAGAGCTGTGCTGAGAGACATATACAGAGAAATGCTGGAGAGTCTGG 885  
Db 241 AspglyAlaGluThrLeuOysLeuGluAspIleTyrThrGluAsnValLeuGluValTyr 260  
QY 886 GCAGATGTGGCAATGGCTGACCCCGCAGAAAGCCAGCCAGCCCTGGGCTGAGAGAG 945  
Db 261 AlaAspValGlyMetCalaGlyProProGlnLysSerProAlaThrLeuGluGluGlu 280  
QY 946 CTCTTCAGCAACCCCTGGCCACCTCATATGACATGCGACACTGCTGCTGGTGGTGGTGA 1005  
Db 281 LeuPheSerThrProGlyHisLeuAsnAspAlaAspThrValLeuValGlyGlu 300  
QY 1006 GCGGAGATGGCAAGACACCTCTCTGACAGCGCTGCACTGCTGTGGGCTGACAGGCA 1065  
Db 301 AlaGlySerGlyLysSerThrLeuGlnArgGluHisLeuThrPalaAlaGlyGln 320  
QY 1066 GACTTCAGAGAAATTTCTTTGTCTTCCATTCAAGTGGCCGCAAGCTGCAATGGCC 1125  
Db 321 AspPheGlnGluPheLeuPheValPheProPheSerCysArgGlnLeuGlnCysMetCala 340  
QY 1126 AAACCACTCTCTGTGGGAGCTACTCTTTGACGACTGCTGTGGCTGATGTGGTCAA 1185  
Db 341 LysProLeuSerValArgThrLeuLeuPheGluHisCysCysTrpProAspValGlyGln 360  
QY 1186 GAAAGACATCTTCCAGTACTCTTGAACACCTGACCCGTGTCTGTAACTTTATGGC 1245  
Db 361 GluAspIlePheGlnLeuLeuLeuAspHisProAspArgValLeuLeuThrPheAspGly 380  
QY 1246 TTGACAGATTCAAGTCAAGGTTCAAGGATGTGAACGCACTGCTCCCGCAGACCCG 1305  
Db 381 PheAspGluPheLysPheArgPheThrAspArgGluArgHisCysSerProThrAspPro 400  
QY 1306 ACCTCTGTCCAGACCTGTCTTCAACCTTCTGACAGGCAACCTGCTGAGAAATGCCGC 1365  
Db 401 ThrSerValGlnThrLeuLeuPheAsnLeuGlnGlyAsnLeuLeuLysAsnAlaArg 420  
QY 1366 AAGGTGCTGACAGACGCTCCGCGCGCTGTGTGGGCTTCTCAAGAAATACATCCGCAC 1425  
Db 421 LysValValThrSerArgProAlaAlaValSerAlaPheLeuArgLysTyrIleArgThr 440  
QY 1426 GAGTTCAACTCAAGGGCTTCTGTGACAGAGGATGAGCTGTACTGAGGAGAGCCCAT 1485  
Db 441 GluPheAsnLeuLysGlyPheSerGlnGlnGlyIleGluLeuTyrLeuArgLysArgHis 460  
QY 1486 CATGAGCCCGGGGTGGCGACCGCTCATCCGCTGCTCCAAAGACCTCAAGCCCTGAC 1545  
Db 461 HisGluProGlyValAlaAspArgLeuLeuLeuArgLeuLeuGlnGlnThrSerAlaLeuHis 480  
QY 1546 GGTGTGTGCACTGCTGCTCTCATGATGTGTCAAAATGCCACAGAGAACTGTTG 1605  
Db 481 GlyLeuOysHisLeuProValPheSerTrpMetValSerLysCysHisGlnGluLeuLeu 500

QY 1606 CTGCAGAGAGGGGGGGTCCCCAAAGACACTACAGATATGTTACCTGCTGATTTGCACGAT 1665  
Db 501 LeuGlnGluGlyGlySerProLysThrThrThAspMetCysLeuLeuIleLeuGlnHis 520  
QY 1666 TTTTGTGCTGATGCCACCCCGCCAACTGAGTCTCCCAAGGTGTGGAGCCCACTCTTCT 1725  
Db 521 PheLeuLeuHisAlaThrProProAspSerAlaSerGlnGluLeuTyrProSerLeuLeu 540  
QY 1726 CCGGCGCGCTCCCAACCTCTGCACCTGGGCAAGTGGCTGTGGGGGCTGGGCAATG 1785  
Db 541 ArgGlyArgLeuProThrLeuLeuHisLeuGluArgGluAlaLeuTrpGlyLeuGlyMet 560  
QY 1786 TGGTCTACGTGTCTCAAGCCAGAGCTCAAGGCAAGAGTCAAGCTGATGACAT 1845  
Db 561 CysCysTyrValPheSerAlaGlnGlnLeuGlnAlaAlaGlnValSerProAspAspIle 580  
QY 1846 TCTCTTGGCTTCTGTGGCGCAAGGTGTGTGGCAGAGATAGCGGCGCCCTGGAA 1905  
Db 581 SerLeuGlyPheLeuValArgAlaLysGlyValAlaProLysSerThrAlaProLeuGlu 600  
QY 1906 TTCTTCAACATCACTTTCAGAGTCTCTTGTCCCGGTTCTTACCTGGCACTCAAGTGTGAT 1965  
Db 601 PheLeuHisIleThrPheGlnCysPhePheAlaAlaPheTyrLeuAlaLeuSerAlaAsp 620  
QY 1966 GTGCCACACCTTTGCTCAGACACCTCTGAAATGTGGCAGCGCAAGCACTCACCAATG 2025  
Db 621 ValProProAlaLeuLeuArgHisLeuPheAsnCysGlyArgProGlyAsnSerProMet 640  
QY 2026 GCCAGGCTCTGCCCCAGATGTGATCCAGGCTCGAGAGGAAAGACAGACAGCTGGCA 2085  
Db 641 AlaArgLeuLeuProThrMetCysIleGlnAlaSerGlnGlyLysAspSerValAla 660  
QY 2086 GCTTTGCTGCAAGAGCCCAAGCCGCAACCTTCAAGTCAAGACAGCTTCTGGCAGGG 2145  
Db 661 AlaLeuLeuGlnLysAlaGluProHisAsnLeuGlnIleThrAlaAlaPheLeuAlaGly 680  
QY 2146 CTGTGTCCCGGAGCACTGGGGCTGCTGGCTGAGTGGCAGACATCTGAGAGGCGCTG 2205  
Db 681 LeuLeuSerArgGluHisTrpGlyLeuLeuAlaGluCysGlnThrSerGluLysAlaLeu 700  
QY 2206 CTCCGGGCGCAAGCCTGTGCCCGCTGGTGTGTGGCCCGCAAGCTCCGAGACACTTCCAC 2265  
Db 701 LeuArgArgGlnAlaCysAlaArgTrpCysLeuAlaArgSerLeuArgGlyHisPheHis 720  
QY 2266 TCCATCCCGGCAAGCTGCAACCGGCTGAGGCGCAAGGCTGCATGCCATGCCCGGTTCA 2325  
Db 721 SerIleProProAlaAlaProGlyAlaAlaLysSerValHisAlaMetProGlyPheIle 740  
QY 2326 TGGCTCATCCGAGCTGTATCAGATGCAGAGAGAGAGGAGTGGTCCGAGAGCTGCAAGT 2385  
Db 741 TrpLeuIleAlaGlySerLeuTyrGluMetGlnGluArgGluAlaArgLysAlaAlaArg 760  
QY 2386 GGGTGAATGTGGGCACTCAAGTGTGCAATTTTGCAAGTGTGGGCCCACTGAGTGTGCT 2445  
Db 761 GlyLeuAsnValGlyHisLeuLysLeuThrPheCysSerValGlyProThrGlnCysAla 780

QY 2446 GCGCTGGCTTGTGCTGCAACCTCCGCGCGCCGCTGCGCTGCACTGGACATACAC 2505  
 Db 781 AAlaLeuAlaPheValLeuGlnHISLeuArgArgProValAlaLeuGlnLeuAspTyrAsn 800  
 QY 2506 TCTGTGGGTGACATTGGCGTGAAGAGCTGGCTGGCTTGGCTGTCTGCAAGGCTCTG 2565  
 Db 801 SerValGlyAspIleGlyValGlnGlnLeuLeuProCysLeuGlnValCysGlyAlaLeu 820  
 QY 2566 TATTGGCGCATTAACAAATATCTGAGACGAGGCAATCTGAAAGCTCAATTGATGCTCTT 2625  
 Db 821 TyrLeuArgAspAsnHISerAspArgGlyLeuCysLeuIleGlnCysAlaLeu 840  
 QY 2626 CACTGCGAACAATTGAGAGATTAGCTCTATTCAACAACAATTGACTGAGCGTGTGCA 2685  
 Db 841 HisCysGlnGlnLeuGlnHISLeuAlaLeuPheAsnHISLeuHISArgGlyCysAla 860  
 QY 2686 CACTGCATGGCTAAAGCTCTTGCATGCGAGCAACAATCTTGGCATTTGAGCTGGGAAT 2745  
 Db 861 HisSerValAlaLeuLeuValAcysArgGlnAspPheLeuAlaLeuArgLeuGlnAsn 880  
 QY 2746 AACTACATCACTGCGCGGAGCCCAAGTGGCTGGCGAGGGGCTCCGAGGCAACATCC 2805  
 Db 881 AsnTyrIleThrAlaAlaGlyAlaGlnValLeuAlaGlnGlyLeuArgGlyAsnHISer 900  
 QY 2806 TTGCAGTTCTGGGATTTGGGGCAACAAGTGGTGAACGAGGGGCCCGAGCCCTGGCT 2865  
 Db 901 LeuGlnPheLeuGlyPheTrpGlyAsnArgValGlyAspGlnGlyAlaGlnAlaLeuAla 920  
 QY 2866 GAGCCCTGGGTGATCACCAAGAGCTTGAAGTGGTGAAGTGGTGGGTGGGGAACAATGGC 2925  
 Db 921 GlnAlaLeuGlnAspHISGlnSerLeuArgTrpLeuSerLeuValGlyAsnAsnIleGly 940  
 QY 2926 AGTGTGGGTGCCAAGCCTTGGCACTGATGCTGGCAAAACAAGTCACTGGTGAAGAATC 2985  
 Db 941 SerValGlyAlaGlnAlaLeuAlaLeuMetLeuAlaLeuValMetLeuGlnGlnLeu 960  
 QY 2986 TGCTGCAAGAGAACCACTCCAGATGAAGTGTATGTTCTCTCCGAGAAAGACTGAAG 3045  
 Db 961 CysLeuGlnGlnAsnHISLeuGlnAspGlnGlyValCysSerLeuAlaGlnGlyLeuLys 980  
 QY 3046 AAAAAATCAAGTTGAAAAATCTGAAGTTGTCATAAATGCACTCACTACCTAAGGGGCA 3105  
 Db 981 LysAsnSerSerLeuLysIleLeuLysLeuSerAspAsnGlyIleThrTyrLeuGlnAla 1000  
 QY 3106 GAAAGCCCTCTGAGGCGCTTGAAGGAAATGACACATCTCGAAGTCTGGCTCCGAGGG 3165  
 Db 1001 GlnAlaLeuLeuGlnAlaLeuGlnArgAspPheIleLeuGlnValIlePheLeuArgGly 1020  
 QY 3166 AACACTTCTCTAGAGAGGTTGACAAGCTCGGCTGCAAGGAGCAACAAGACTTGGCT 3225  
 Db 1021 AsnThrPheSerLeuGlnGlnValAspLysLeuGlnCysArgAspPheArgLeuLeuLeu 1040

Search completed: January 7, 2004, 19:37:47  
 Job time : 384.738 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 7, 2004, 19:49:02 ; Search time 133.985 Seconds

(without alignments)  
 13490.627 Million cell updates/sec

Title: US-10-014-269-1  
 Perfect score: 8270  
 Sequence: 1 gtgacagatccagctcac.....ataactctgtagctcaaac 4465

Scoring table: BLOSUM62

Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 747907 segs, 201509753 residues

Total number of hits satisfying chosen parameters: 1495814

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Command line parameters:

-MODEL=framet\_n2p.model -DEV=wxip

Q:/cgn2\_1/USPRO\_spool\_p/US10014269/runat\_07012004\_175231\_24632/app\_query.fasta\_1  
 .9358  
 -DB=Published\_Applications\_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
 -IOPCla0 -IOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blcosum62  
 -TRANS=human40.cdd -LIST=150 -DOCALLIGN=200 -THR SCORE=pcpt -THR MAX=100  
 -THR MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptio -NORWext -HEAPSIZE=500 -MINLEN=0  
 -MAXLEN=2000000000 -USER=US10014269 GCN 1.1 5.9 runat\_07012004\_175231\_24632  
 -NCPU=6 -ICPU=3 -NO\_MMAP -LARGEJOBERT -NEG\_SCORES=0 -WAIT -DSPLOC=100  
 -LONGJOB -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep:\*\n2: /cgn2\_6/prodata/2/pubpaa/ECT\_NEW\_PUB.pep:\*\n3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep:\*\n4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep:\*\n5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep:\*\n6: /cgn2\_6/prodata/2/pubpaa/ECTUS\_PUBCOMB.pep:\*\n7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep:\*\n8: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep:\*\n9: /cgn2\_6/prodata/2/pubpaa/US09b\_PUBCOMB.pep:\*\n10: /cgn2\_6/prodata/2/pubpaa/US09b\_PUBCOMB3.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5448	65.9	1040	12	US-10-314-506-2	Sequence 2, Appl
2	5448	65.9	1040	14	US-10-014-269-2	Sequence 2, Appl
3	5448	65.9	1040	14	US-10-002-974-2	Sequence 2, Appl
4	5447	65.9	1040	14	US-10-002-974-67	Sequence 67, Appl
5	5445	65.8	1040	14	US-10-002-974-63	Sequence 63, Appl
6	5444	65.8	1040	14	US-10-002-974-69	Sequence 69, Appl
7	5443	65.8	1040	12	US-10-314-506-4	Sequence 4, Appl
8	5443	65.8	1040	14	US-10-014-269-4	Sequence 4, Appl
9	5443	65.8	1040	14	US-10-002-974-4	Sequence 4, Appl
10	5443	65.8	1040	14	US-10-002-974-65	Sequence 65, Appl
11	5440	65.8	1040	12	US-10-314-506-57	Sequence 57, Appl
12	5440	65.8	1040	14	US-10-002-974-59	Sequence 59, Appl
13	5440	65.8	1040	14	US-10-002-974-61	Sequence 61, Appl
14	5440	65.8	1040	14	US-10-002-974-61	Sequence 61, Appl
15	5435	65.7	1040	14	US-10-002-974-87	Sequence 87, Appl
16	5432	65.7	1040	12	US-10-314-506-59	Sequence 59, Appl
17	5432	65.7	1040	12	US-10-314-506-61	Sequence 61, Appl
18	5432	65.7	1040	14	US-10-002-974-85	Sequence 85, Appl
19	5432	65.7	1040	14	US-10-002-974-89	Sequence 89, Appl
20	5331	64.5	1041	12	US-10-240-046-2	Sequence 2, Appl
21	5304	64.1	1013	12	US-10-314-506-3	Sequence 3, Appl
22	5304	64.1	1013	14	US-10-014-269-3	Sequence 3, Appl
23	5304	64.1	1013	14	US-10-002-974-3	Sequence 3, Appl
24	5272	63.7	1007	12	US-10-014-506-34	Sequence 34, Appl
25	5272	63.7	1007	14	US-10-014-269-34	Sequence 34, Appl
26	5272	63.7	1007	14	US-10-002-974-34	Sequence 34, Appl
27	5264	63.7	1007	12	US-10-314-506-55	Sequence 55, Appl
28	5264	63.7	1007	14	US-10-002-974-55	Sequence 55, Appl
29	4994.5	60.4	1009	10	US-09-864-921-107	Sequence 107, App
30	4188	35.8	795	10	US-09-864-921-188	Sequence 188, App
31	2959	35.8	560	10	US-09-864-921-83	Sequence 83, Appl
32	2087	25.2	403	10	US-09-864-921-176	Sequence 176, App
33	1614	19.5	305	12	US-10-014-269-7	Sequence 7, Appl
34	1614	19.5	305	14	US-10-002-974-7	Sequence 7, Appl
35	1614	19.5	305	14	US-10-002-974-7	Sequence 7, Appl
36	1564	18.9	296	10	US-09-864-921-174	Sequence 174, App
37	1219.5	14.7	953	9	US-09-728-721-43	Sequence 43, Appl
38	1219.5	14.7	953	14	US-10-118-584-43	Sequence 43, Appl
39	1219.5	14.7	953	15	US-10-295-981-43	Sequence 43, Appl
40	1209.5	14.6	953	9	US-09-728-721-8	Sequence 8, Appl
41	1209.5	14.6	953	12	US-10-028-374-4	Sequence 4, Appl
42	1209.5	14.6	953	12	US-10-183-770-4	Sequence 4, Appl
43	1209.5	14.6	953	14	US-10-105-931-8	Sequence 8, Appl
44	1209.5	14.6	953	14	US-10-118-984-8	Sequence 8, Appl
45	1209.5	14.6	953	15	US-10-013-477-12	Sequence 12, Appl
46	1209.5	14.6	953	15	US-10-295-981-8	Sequence 8, Appl
47	1201	12.1	705	15	US-10-013-477-19	Sequence 19, Appl
48	1001	12.1	705	15	US-10-013-477-19	Sequence 19, Appl
49	1001	12.1	719	15	US-10-106-698-5363	Sequence 6363, Ap
50	858	10.4	166	9	US-09-864-761-35000	Sequence 35000, A
51	840	10.2	155	9	US-09-864-761-41737	Sequence 41737, A
52	820	9.9	146	9	US-09-864-761-41738	Sequence 41738, A
53	652.5	7.9	180	10	US-09-864-921-189	Sequence 189, App
54	596	7.2	1121	12	US-10-161-927-2	Sequence 2, Appl
55	593	7.2	140	10	US-09-864-921-191	Sequence 191, App
56	588	7.1	139	10	US-09-864-921-193	Sequence 193, App
57	532.5	6.4	1027	12	US-10-407-866-68	Sequence 68, Appl
58	531	6.4	1061	15	US-10-066-521-18	Sequence 18, Appl
59	530.5	6.4	490	9	US-09-728-721-26	Sequence 26, Appl
60	530.5	6.4	490	14	US-10-105-931-26	Sequence 26, Appl
61	530.5	6.4	490	14	US-10-118-984-26	Sequence 26, Appl
62	530.5	6.4	490	15	US-10-295-981-26	Sequence 26, Appl
63	530.5	6.4	490	15	US-10-295-981-26	Sequence 26, Appl
64	530.5	6.4	1035	11	US-09-965-621-24	Sequence 24, Appl
65	522	6.3	97	12	US-10-314-506-24	Sequence 24, Appl
66	522	6.3	97	14	US-10-014-269-5	Sequence 5, Appl
67	522	6.3	97	14	US-10-002-974-5	Sequence 5, Appl
68	509.5	6.2	507	12	US-10-401-194-3	Sequence 3, Appl
69	497	6.0	1034	12	US-10-028-374-18	Sequence 18, Appl
70	497	6.0	1034	12	US-10-132-967-18	Sequence 18, Appl
71	497	6.0	1034	12	US-10-183-770-18	Sequence 18, Appl
72	497	6.0	1034	14	US-10-127-516-5	Sequence 5, Appl
73	497	6.0	1034	14	US-10-027-629-5	Sequence 5, Appl
74	485	5.9	94	12	US-10-314-506-6	Sequence 6, Appl
75	485	5.9	94	14	US-10-014-269-6	Sequence 6, Appl
76	485	5.9	94	14	US-10-002-974-6	Sequence 6, Appl
77	433	5.2	1106	10	US-09-064-199-17	Sequence 17, Appl
78	433	5.2	1130	10	US-09-064-199-18	Sequence 18, Appl
79	433	5.2	1130	14	US-10-104-599-6	Sequence 6, Appl
80	433	5.2	1207	10	US-09-064-199-16	Sequence 16, Appl
81	430.5	5.2	1016	10	US-09-866-224-19	Sequence 19, Appl
82	429.5	5.2	896	15	US-10-066-521-22	Sequence 22, Appl
83	425	5.1	320	12	US-10-314-506-30	Sequence 30, Appl
84	425	5.1	320	14	US-10-014-269-30	Sequence 30, Appl
85	425	5.1	320	14	US-10-002-974-30	Sequence 30, Appl
86	418	5.1	81	10	US-09-864-921-172	Sequence 172, App
87	415	5.0	1454	10	US-09-388-721-10	Sequence 10, Appl
88	414.5	5.0	1429	12	US-10-028-374-3	Sequence 3, Appl
89	414.5	5.0	1429	12	US-10-183-770-3	Sequence 3, Appl
90	412.5	5.0	1399	10	US-09-388-221-4	Sequence 4, Appl
91	412.5	5.0	1443	10	US-09-388-221-6	Sequence 6, Appl
92	411.5	5.0	1424	10	US-09-388-221-12	Sequence 12, Appl
93	411	5.0	830	10	US-09-064-199-19	Sequence 19, Appl
94	410.5	5.0	1429	10	US-09-996-671-2	Sequence 2, Appl
95	410.5	5.0	1429	10	US-09-331-071-2	Sequence 2, Appl
96	410.5	5.0	1429	12	US-10-028-374-15	Sequence 15, Appl



97 410.5 5.0 1429 12 US-10-183-770-15  
 98 410.5 5.0 1429 15 US-10-028-392-11  
 99 410.5 5.0 1473 10 US-09-388-221-2  
 100 408 4.9 76 10 US-09-864-921-170  
 101 399.5 4.8 919 12 US-10-094-749-2718  
 102 396 4.8 858 9 US-09-848-035-8  
 103 396 4.8 858 10 US-09-986-224-8  
 104 396 4.8 994 11 US-09-965-621-16  
 105 396 4.8 994 12 US-10-407-866-16  
 106 396 4.8 994 15 US-10-066-521-24  
 107 395 4.8 952 12 US-10-407-866-70  
 108 390.5 4.7 1049 12 US-10-233-663-42  
 109 385 4.7 980 9 US-09-848-035-13  
 110 385 4.7 980 10 US-09-986-224-13  
 111 381 4.6 772 12 US-10-094-749-2866  
 112 379.5 4.6 1062 12 US-10-239-663-43  
 113 379 4.6 635 12 US-10-407-866-90  
 114 375 4.5 1033 12 US-10-132-967-2  
 115 375 4.5 1033 14 US-10-127-516-2  
 116 375 4.5 1033 14 US-10-027-629-2  
 117 366 4.4 947 11 US-09-965-621-18  
 118 366 4.4 947 12 US-10-407-866-18  
 119 365.5 4.4 1111 15 US-10-216-645-5  
 120 364 4.4 1344 15 US-10-066-521-6  
 121 362 4.4 902 12 US-10-407-866-66  
 122 357 4.3 1143 15 US-10-216-645-4  
 123 357 4.3 1162 15 US-10-216-645-2  
 124 352.5 4.3 674 12 US-10-407-866-89  
 125 352.5 4.3 719 11 US-09-965-621-59  
 126 352.5 4.3 719 12 US-10-407-866-59  
 127 350 4.2 612 12 US-10-407-866-88  
 128 339.5 4.1 655 12 US-10-296-539-2  
 129 339.5 4.1 655 12 US-10-407-866-84  
 130 339.5 4.1 655 15 US-10-066-521-13  
 131 339.5 4.1 720 12 US-10-342-331-4  
 132 339 4.1 590 12 US-10-407-866-94  
 133 338 4.1 720 12 US-10-342-331-4  
 134 337 4.1 1466 12 US-10-301-822-33  
 135 337 4.1 1466 15 US-10-177-293-68  
 136 331 4.0 584 12 US-10-029-386-32446  
 137 329.5 4.0 821 9 US-09-848-035-2  
 138 329.5 4.0 821 10 US-09-986-224-2  
 139 329 4.0 200 9 US-09-728-721-11  
 140 329 4.0 200 14 US-10-105-931-11  
 141 329 4.0 200 14 US-10-118-984-11  
 142 329 4.0 200 15 US-10-285-981-11  
 143 323.5 3.9 321 12 US-10-407-866-116  
 144 323.5 3.9 960 12 US-10-342-331-5  
 145 322.5 3.9 1078 15 US-10-058-124-21  
 146 321.5 3.9 659 12 US-10-407-866-91  
 147 319.5 3.9 875 9 US-09-848-035-16  
 148 319.5 3.9 875 10 US-09-986-224-16  
 149 318 3.9 4123 15 US-10-213-509-5  
 150 315 3.8 960 12 US-10-342-331-5

# ALIGNMENTS

Sequence 15, Appl  
 Sequence 11, Appl  
 Sequence 2, Appl  
 Sequence 170, App  
 Sequence 2718, Ap  
 Sequence 8, Appl  
 Sequence 8, Appl  
 Sequence 16, Appl  
 Sequence 16, Appl  
 Sequence 24, Appl  
 Sequence 42, Appl  
 Sequence 13, Appl  
 Sequence 13, Appl  
 Sequence 2566, Ap  
 Sequence 43, Appl  
 Sequence 90, Appl  
 Sequence 2, Appl  
 Sequence 2, Appl  
 Sequence 2, Appl  
 Sequence 18, Appl  
 Sequence 18, Appl  
 Sequence 5, Appl  
 Sequence 6, Appl  
 Sequence 66, Appl  
 Sequence 4, Appl  
 Sequence 89, Appl  
 Sequence 59, Appl  
 Sequence 59, Appl  
 Sequence 88, Appl  
 Sequence 2, Appl  
 Sequence 84, Appl  
 Sequence 4, Appl  
 Sequence 94, Appl  
 Sequence 4, Appl  
 Sequence 68, Appl  
 Sequence 33, Appl  
 Sequence 3246, A  
 Sequence 2, Appl  
 Sequence 2, Appl  
 Sequence 11, Appl  
 Sequence 11, Appl  
 Sequence 11, Appl  
 Sequence 116, Appl  
 Sequence 21, Appl  
 Sequence 91, Appl  
 Sequence 16, Appl  
 Sequence 16, Appl  
 Sequence 5, Appl  
 Sequence 5, Appl

RESULT 1  
 US-10-314-506-2  
 / Sequence 2, Application US/10314506  
 / Publication No. US20030175762N1  
 / GENERAL INFORMATION:  
 / APPLICANT: Inohara, Naohiro  
 / APPLICANT: Ogura, Yasunori  
 / TITLE OF INVENTION: Modulators of NOD2 signaling  
 / FILE REFERENCE: UM-06984  
 / CURRENT APPLICATION NUMBER: US/10/314,506  
 / PRIOR FILING DATE: 2002-12-09  
 / PRIOR APPLICATION NUMBER: 10,014,269  
 / PRIOR FILING DATE: 2001-10-26  
 / PRIOR APPLICATION NUMBER: 60/244,299  
 / NUMBER OF SEQ ID NOS: 62  
 / SOFTWARE: PatentIn version 3.1  
 / SEQ ID NO 2  
 / LENGTH: 1040  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 US-10-314-506-2  
 Alignment Scores:  
 Pred. No.: 0  
 Score: 5448.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 65.88%  
 Gaps: 0  
 Length: 1040  
 Matches: 1040  
 Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0  
 US-10-014-269-1 (1-4485) x US-10-314-506-2 (1-1040)  
 QY 106 ATGGGGGAAGAAGGGGTTCACCTCTCACGATGAGAGCAAGCAAGTCTCTC 165  
 |||||||  
 Db 1 MetGlyuLugLgLySerLAserHisAspGluGluGluArgLAserValLeuLeu 20  
 QY 166 GGACATCTCCGGGTGTGAATGTGCTCCGACGAGAGGCTTTCAAGCAAGAGCGCA 225  
 |||||||  
 Db 21 GlyHisSerProGlyCySgluWetCySerGlnGlnAlaHedLAlaGlnArgSerLIn 40  
 QY 226 CTGGTCAAGCTCTGGTCTCAGAGGTCCTTGGAAGGCTTCGAGAGTCTCGAGCTGGCTG 285  
 |||||||  
 Db 41 LeuValGluLeuLeuValSerGlySerLeuGluGlyPheGluSerValLeuAspTrpLeu 60  
 QY 286 CTGTCTGGGAGGTCCTCTCTCGGAGAGCTACAGAGGCTTCACCTCTGGGCGAGGCT 345  
 |||||||  
 Db 61 LeuSerTrpGluValLeuSerTrpGluAspTrpGluGlyPheHisLeuLeuGlyGlnPro 80  
 QY 346 CTCTCCCACTTGGCCAGGCGCTTCTGGAACAGCTCGAATTAAGGTAATTGGGCTGT 405  
 |||||||  
 Db 81 LeuSerHisLeuAlaArgArgLeuLeuAspTrpValTrpAsnLysGlyTrpTrpAlaCys 100  
 QY 406 CGAAGCTCATCGGGGCTGCGGAGAGAGGCTCGCAAGGCGCAAGCAATCCCCAAGCTGAT 465  
 |||||||

Db 101 GlnLysLeuIleAlaAlaLeuGlnLysAlaGlnLysAspSerGlnSerProLysLeuHis 120  
QY 466 GGCTGCTGGACCCCACTCGCTCCACCAAGCCGAGACCTGCAAGACTGACCGGACGCC 525  
Db 121 GlyCysTrpAspProHisSerLeuHisAspProAlaArgAspLeuGlnSerHisArgProAla 140  
QY 526 ATGTGAGAGGCTCCACAGCACTGTGGAGAAACATGCTGAGACTGGCATGGAGAGGGGGT 585  
Db 141 IleValArgAlaGlnLeuHisSerHisValGlnLysMetLeuAspLeuAlaTrpGlnArgGly 160  
QY 586 TTGCTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 645  
Db 161 PheValSerGlnTyrGlnCysAspGlnIleArgLeuProIlePheThrProSerGlnArg 180  
QY 646 GCAAGAGGCTGCTGATCTTGCCACGGTGAAGAAACCAATGGATGGCTGCTCTCTCA 705  
Db 181 AlaArgAlaLeuLeuAspLeuAlaThrValLysAlaAsnGlyLeuAlaAlaPheLeu 200  
QY 706 CAACATGTTCAAGAAATTACAGTCCCATGGCCCTGCTTTGAAAGCTGCAATGCAAG 765  
Db 201 GlnHisValGlnGlnLeuProValProLeuAlaLeuProLeuGlnAlaAlaThrCysLys 220  
QY 766 AAGTATATGGCAAGCTGAGAGCAAGGCTGCTCTCACTGCTCTCTCTCTCTCACTACT 825  
Db 221 LysTyrMetValLysLeuArgThrThrValSerAlaGlnSerArgPheLeuSerThrTyr 240  
QY 826 GATGAGAGAGAGAGCTGCTGCTGAGAGCATATACAGAAATGCTGAGAGTCTGG 885  
Db 241 AspGlyAlaGlnThrLeuCysLeuGlnLysAlaTyrThrGlnMetAlaLeuGlnValTrp 260  
QY 886 GCAAGATGTGGCACTGGCTGACCCCGCAGAAAGCCCAAGCCCTGGGCTGAGAGAG 945  
Db 261 AlaAspValGlyMetAlaGlyProProGlnLysSerProAlaThrLeuGlyLeuGln 280  
QY 946 CTCTTCAGCAAGCCCTGGCCACTCAATGACATGCGGACATGCTGCTGGTGGGTGAG 1005  
Db 281 LeuPheSerThrProLysHisLeuLysAspAspAlaAspThrValLeuValAlaGlyGln 300  
QY 1006 GCGGCACTGGCAAGAGACGCTCTGCAAGCGGCTGCACCTTGTGTGGGCTGCAAGGCA 1065  
Db 301 AlaGlySerGlyLysSerThrLeuLeuGlnArgLeuHisLeuLeuTrpAlaAlaGlyGln 320  
QY 1066 GACTTCAGAGAAATTCTTGTCTTCCATTGACGCTGCGGCAAGCTGCTGCTGCTGCTGCT 1125  
Db 321 AspPheGlnGlnPheLeuPheValPheProPheSerCysArgGlnLeuGlnCysMetAla 340  
QY 1126 AAACCACTCTGTGCGGACTTACTCTTTGAGCACTGCTGTTGGCTGATGTTGGTCAA 1185  
Db 341 LysProLeuSerValArgThrLeuLeuPheGlnHisCysCysTrpProAspValGlyGln 360  
QY 1186 GAAGACATCTTCAGTTACTCTTGACCAAGCTGACCGTGTCTCTTTAACTTTAGTGGC 1245  
Db 361 GlnAspIlePheGlnLeuLeuAspHisProAspAlaGlyValLeuLeuThrPheAspGly 380  
QY 1246 TTGACAGATTCAAGTTCAGGTTCAAGGATGATGAAGCACTGCTCCCGACCAAGCCGCC 1305  
Db 381 PheAspGlnPheLysPheArgPheThrAspArgGlnLysGlnHisCysSerProThrAspPro 400

QY 1306 AACTCTGTCAGAACCTGCTCTTCAACCTTCTGCAAGGCAACCTGCTGAAGATGCCGCC 1365  
Db 401 ThrSerValGlnThrLeuLeuPheLeuLeuGlnGlnLysLeuLeuLysAspAlaArg 420  
QY 1366 AAGTGTGACCAAGCCGTCCGGGCGGTGTGTGCGGCTCTCTCAAGAAATGATCCCGCAC 1425  
Db 421 LysValValThrSerArgProAlaAlaValSerAlaPheLeuArgLysTyrIleArgThr 440  
QY 1426 GAGTTCACACTCAAGGCGCTTCTCTGACAGGCGCATCGAGCTGATGAGAGAGCGCAT 1485  
Db 441 GlnPheAsnLeuLysGlyPheSerGlnGlnGlyTyrIleGlnLeuTyrLeuArgLysArgHis 460  
QY 1486 CATTAGCCCCGGGTGGGAGACCGCTCATCGCGCTGCTCCAAAGAACTCAGCCCTGCAC 1545  
Db 461 HisGlnProGlyValAlaAspArgLeuIleArgLeuLeuGlnGlnThrSerAlaLeuHis 480  
QY 1546 GATTGTGCAACCTGCTCTCTCATGATGATGTCAAATGCCAAGCAAGAACTGTTG 1605  
Db 481 GlyLeuCysHisLeuProValPheSerTrpMetValSerLysCysHisGlnLeuLeu 500  
QY 1606 CTGCAAGAGGGGGGCTCCCAAGCAACTACAGATATGTAAGTCTGATTTCTGACGAT 1665  
Db 501 LeuGlnGlnGlyLysSerProLysThrThrAspMetTyrLeuLeuIleLeuGlnHis 520  
QY 1666 TTCTGTGATGATGCAACCCCGCCAGACTGCTCCCAAGGCTGGAGCCCAAGTCTCT 1725  
Db 521 PheLeuLeuHisAlaThrProProAspSerAlaSerGlnGlyLeuGlyProSerLeuLeu 540  
QY 1726 CCGGCGCGGCTCCCAAGCTCTGCAAGCTGGGCAAGTGGCTGTGGGGGCTGGGCAAG 1785  
Db 541 ArgGlyArgLeuProThrLeuLeuHisLeuGlyArgLeuAlaLeuTrpGlyLeuGlyMet 560  
QY 1786 TGTGCTACGTTGCTCTCAAGCAAGCTGTCAGAGCAAGCTCAGCCCTGATGACATT 1845  
Db 561 CysCysTyrValPheSerAlaGlnGlnLeuGlnAlaAlaGlnValSerProAspAspIle 580  
QY 1846 TCTCTTGGCTTCTGCTGCTGCTGCAAGGCTGCTGCGCAGAGTACGCGCCCTGGAA 1905  
Db 581 SerLeuGlyPheLeuValArgAlaLysGlyValValProGlySerThrAlaProLeuGln 600  
QY 1906 TTCTTCACATCACTTTCAGTGCTCTTTCGCGGCTTCACTGGCACTCAAGTGTGAT 1965  
Db 601 PheLeuHisIleThrPheGlnCysPheSerAlaAlaPheTyrLeuAlaLeuSerAlaAsp 620  
QY 1966 GTGCCACCACTTTGCTCAGACACCTCTCAATGTGGCAGGCAAGCACTCAACCATG 2025  
Db 621 ValProProAlaLeuLeuArgHisLeuPheAsnGlyArgProGlyAsnSerProIle 640  
QY 2026 GCAAGGCTCTGCGCAAGATGTCATCCAGGCTCGAGAGGAAAGAACAGACGACCTGGCA 2085  
Db 641 AlaArgLeuLeuProThrMetCysIleGlnAlaSerGlnGlyLysAspSerValAla 660  
QY 2086 GCTTGTGTCAGAGGCGCAAGGCGCAAGCACTTCAAGTCAAGGAGCTTCTGCGCAGG 2145  
Db 661 AlaLeuLeuHisLysAlaGlnProHisAsnLeuGlnIleThrAlaAlaPheLeuAlaGly 680

QY 2146 CTGTGTCCCGGAGACATGCGGGCTGTGTGCTGAGTGCACACATCTGAGAGGCTCTG 2205  
 Db |||||  
 681 LeuLeuSerPheArgIleHisTrpGlyLeuLeuAlaGluCysGlnHisSerGlnLysAlaLeu 700  
 QY 2206 CTCGGCGCCAGAGCCTGTGCCGCGCTGGTGTGTGGCCGCGAGCTCCGACAGCACTTCCAC 2265  
 Db |||||  
 701 LeuArgArgGlnAlaCysAlaArgTrpCysLeuAlaArgSerLeuAlaArgLysHisPheHis 720  
 QY 2266 TCACATCCCGCAGCTGCACCGGGGTGAAGCCAGACCTGCATGCCATGCCCGGGTTCATC 2325  
 Db |||||  
 721 SerIleProPheAlaAlaProGlyValAlaLysSerValHisAlaMetProGlyPheIle 740  
 QY 2326 TGGCTCATCCGGAGGCTGTACAGATGCAGAGAGAGCGGCTGGCTCGAGAGAGCTGCACTG 2385  
 Db |||||  
 741 TrpLeuIleArgSerLeuTyrGlnMetGlnGluGlnArgLeuAlaArgLysAlaAlaArg 760  
 QY 2386 GGCTGAAATGCTGGGACACTCAAGTGAACATTTTGCAGTGTGGGCCCTGACTGATGCTGCT 2445  
 Db |||||  
 761 GlyLeuAsnValGlyHisLeuLysLeuThrPheCysSerValGlyProThrGlnCysAla 780  
 QY 2446 GGCTTGGCTTGTGTGCAACACCTCCGCGGGCCGTGGCCCTGCACTGGAGACTACAC 2505  
 Db |||||  
 781 AlaLeuAlaPheValLeuGlnHisLeuArgArgProValAlaLeuGlnLeuAspTyrAsn 800  
 QY 2506 TCTGTGGGTGACATTTGGCTGGAGCACTGCTGCTTGGCTTGGTGTGCAAGGCTCTG 2565  
 Db |||||  
 801 SerValGlyAspIleGlyValGlnGlnLeuProCysLeuGlyValCysLysAlaLeu 820  
 QY 2566 TATTGGCGCAATTAACAAATCTCAGACCGAGGCATCTGCAAGCTCATTGATGATGCTT 2625  
 Db |||||  
 821 TyrLeuArgPhePheAsnHisIleSerPheArgGlyIleCysLysLeuIleGlnCysAlaLeu 840  
 QY 2626 CACTGCGACCAATTGCAGAACTTACTTATTCACCAACAAATTGACTGACGCTGTGCA 2685  
 Db |||||  
 841 HisCysGlnGlnLeuGlnLysLeuAlaLeuPheAsnLysLeuThrAspGlyCysAla 860  
 QY 2686 CACTGCATGGCTAAGCTCCTTGCATGCAGCGCAAACTTCTTGGCATTGAAGCTGGGGAAT 2745  
 Db |||||  
 861 HisSerMetAlaLysLeuLeuAlaCysArgGlnAsnPheLeuAlaLeuArgLeuGlyAsn 880  
 QY 2746 AACTACATCACTGCCGCGGAGACCAAGTGTGGCTGGCCGAGGGGCTCCGAGGCAACACTCC 2805  
 Db |||||  
 881 AsnTyrIleThrAlaAlaGlyAlaGlnValLeuAlaGlnGlyLeuArgGlyAsnThrSer 900  
 QY 2806 TTGCAGTCTCGGATTTCTGGGGCAACAAGTGGGTACAGAGGGGCCCGACGCTGGCT 2865  
 Db |||||  
 901 LeuGlnPheLeuGlyPheTrpGlyAsnArgValGlyAspGlnGlyAlaGlnAlaLeuAla 920  
 QY 2866 GAAGCTTGGGTGATCACAAGAGCTTGAAGTGGCTCAAGCTGGTGGGGAAACAACATTGGC 2925  
 Db |||||  
 921 GluAlaLeuGlyAspHisGlnSerLeuArgTrpLeuSerLeuValGlyAsnAsnIleGly 940  
 QY 2926 AGTGTGGGTGCCAAGCTTGGCACTGATGCTGGCAAAAGACGTCATGCTAGAGAACTC 2985  
 Db |||||  
 941 SerValGlyAlaGlnAlaLeuAlaLeuMetLeuAlaLysAsnValMetLeuGlnGlnLeu 960  
 QY 2986 TGGCTGGAAGAGAAACATCTCCAGATGAAAGGTGTATGTTCTCCGACAGAGACTGAAG 3045

Db |||||  
 961 CysLeuGlnGlnLysHisLeuGlnAspGlnGlyValCysSerLeuAlaGlnGlyLeuLys 980  
 QY 3046 AAAAATTCAATTGAAAAATCTGAAGTGTCCATTAAGTGCATCACTACCTAGCTAGGGGCA 3105  
 Db |||||  
 981 LysAsnSerSerLeuLysIleLeuLysLeuSerPheAsnCysIleThrTyrLeuGlyAla 1000  
 QY 3106 GAAGCCTTCTGCAAGGCTTGAAGGAATGACACCATCTGGAAGTCTGGCTCCGAGGG 3165  
 Db |||||  
 1001 GluAlaLeuLeuGlnAlaLeuGlnArgAsnAspThrIleLeuGlnValTrpLeuArgGly 1020  
 QY 3166 AACACTTCTCTAGAGAGAGTGTGAAGAGCTGGAGCTGGAGGAGACACAGACTGTGCT 3225  
 Db |||||  
 1021 AsnThrPheSerLeuGlnGlnLysValAspLysLeuGlyCysArgPhePheArgLeuLeu 1040

Search completed: January 7, 2004, 20:49:14  
 Job time : 432.985 secs

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 7, 2004, 19:24:31 ; Search time 65.4927 seconds

(without alignments)  
13171.439 Million cell updates/sec

Title: US-10-014-269-1  
Perfect score: 8270  
Sequence: 1 gtagacagatccaggtccac.....ataacgtgttagtcaaac 4485

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Command line parameters:  
-MODE=frame\_plus\_n2p.model -DEV=xlp

Q=/cgn2\_1/USPTO\_spool\_p/US10014269/rnat\_07012004\_175229\_24535/app\_query.fasta\_1  
-DB=PIR\_76 -QMT=fastan -SUFFIX=pr -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=150  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50 -MODE=LOCAL  
-OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us10014269/cgn\_1\_1\_224/grunat\_07012004\_175229\_24535 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORE=0 -WAIT -DSPLICER=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : PIR\_76:++  
1: pir1:++  
2: pir2:++  
3: pir3:++  
4: pir4:++

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB ID	Description
		Match	Length		
1	433	5.2	1130	2 A48843	MHC class II trans
2	393.5	4.8	1192	2 T17255	hypothetical prote
3	365.5	4.4	1111	2 A59000	matern protein (imp
4	337	4.1	1466	1 CGHUL	collagen alpha 1(I
5	320	3.9	1453	2 S21626	collagen alpha 1(I
6	314.5	3.8	1373	1 A43281	collagen alpha 1(I
7	314	3.8	1464	2 S59856	collagen alpha 1(I
8	312.5	3.8	1464	2 S59856	collagen alpha 1(I
9	312	3.8	784	2 U00317	hypothetical 82k P
10	311	3.8	1464	2 U00317	collagen alpha 1(I
11	308.5	3.8	660	1 Q0E83	BHLFI protein - hu
12	308	3.8	2944	2 A54849	collagen alpha 1(I
13	303	3.7	1466	1 CGHUL	collagen alpha 1(I
14	301.5	3.6	1691	1 S22917	collagen alpha 5(I
15	301	3.7	1496	1 CGHUL	collagen alpha 2(I
16	300.5	3.7	1049	1 CGB075	collagen alpha 1(I
17	294.5	3.6	1049	1 CGB075	collagen alpha 1(I
18	294	3.6	779	1 CGB015	collagen alpha 1(I
19	292.5	3.6	1414	1 S23809	collagen alpha 2(I
20	291.5	3.5	1487	1 CGHUC	collagen alpha 2(I
21	290	3.5	1690	1 CGHUL	collagen alpha 4(I
22	289	3.5	886	2 I50694	collagen alpha 1(I
23	287.5	3.5	1366	1 CGHUL	collagen alpha 1(I
24	287.5	3.5	1670	1 CGHUL	collagen alpha 3(I
25	287	3.5	886	2 I50694	collagen alpha 1(I
26	286.5	3.5	1418	2 T45467	collagen alpha 1(I
27	286.5	3.5	2715	2 T13049	eyelid - fruit fly
28	284	3.5	671	1 CGBR15	collagen alpha 1(I
29	282	3.4	2944	2 A54849	collagen alpha 1(I
30	281.5	3.4	1806	1 CGHUL	collagen alpha 1(I
31	281	3.4	1690	1 CGHUL	collagen alpha 1(I
32	280.5	3.4	1414	1 S23809	collagen alpha 2(I
33	279.5	3.4	1414	1 CGHUL	collagen alpha 2(I
34	276	3.3	1453	2 S21626	collagen alpha 1(I
35	275	3.3	1042	1 CGHUL	collagen alpha 1(I
36	274	3.3	1497	2 T45467	procollagen type V
37	271.5	3.3	1418	2 T45467	collagen alpha 1(I
38	271.5	3.3	1419	2 A41182	collagen alpha 1(I
39	271.5	3.3	1487	2 B41182	collagen alpha 1(I
40	271	3.3	456	2 A31857	ribonuclease inhib
41	270	3.3	1419	2 A41182	collagen alpha 1(I
42	270	3.3	1487	2 B41182	collagen alpha 1(I
43	264.5	3.2	1838	1 CGHUL	collagen alpha 1(I
44	264	3.2	1985	2 T31611	hypothetical prote
45	263.5	3.2	1496	1 CGHUL	collagen alpha 2(I
46	263	3.2	1487	1 CGHUL	collagen alpha 1(I
47	263	3.2	1669	1 CGHUL	collagen alpha 1(I
48	262.5	3.2	1464	1 CGHUL	collagen alpha 1(I
49	262	3.2	1042	1 CGHUL	collagen alpha 1(I
50	258.5	3.1	1492	2 A40333	collagen alpha 1(I
51	258.5	3.1	1763	2 S16366	collagen alpha 2(I
52	258.5	3.1	1806	1 CGHUL	collagen alpha 1(I
53	258.5	3.1	1843	2 S18803	collagen alpha 1(I

54	258	3.1	1492	2	A40333	collagen alpha 1(I)
55	257	3.1	1779	1	CG8035	collagen alpha 1(I)
56	257	3.1	1843	2	S18603	collagen alpha 1(I)
57	256.5	3.1	1669	1	CGH048	collagen alpha 1(I)
58	256.5	3.1	1838	1	CGH01V	collagen alpha 1(I)
59	255.5	3.1	888	2	S28791	collagen alpha 1(I)
60	255.5	3.1	1691	1	S22971	collagen alpha 1(I)
61	252.5	3.1	2715	2	T13049	eyelid - fruit fly
62	252	3.0	660	1	Q0853	BHLH protein - hu
63	252	3.1	1603	2	S23810	collagen alpha 1(I)
64	251.5	3.0	1486	1	B40333	collagen alpha 1(I)
65	250.5	3.1	1486	1	B40333	collagen alpha 1(I)
66	250	3.0	1373	1	A43291	collagen alpha 1(I)
67	248.5	3.0	1603	2	S23810	collagen alpha 1(I)
68	247.5	3.0	1670	1	CGH03B	collagen alpha 1(I)
69	246.5	3.0	920	2	B34433	collagen alpha 1(I)
70	244.5	3.0	1366	1	CGH02S	collagen alpha 1(I)
71	243	2.9	1357	2	T29265	hypothetical prote
72	242	2.9	4957	2	T03455	ALR protein - huma
73	242	2.9	5262	2	T03455	collagen alpha 1(I)
74	241	2.9	744	1	A34246	collagen alpha 1(I)
75	241	2.9	744	1	S23298	collagen alpha 1(I)
76	240.5	2.9	964	1	CGH02S	collagen alpha 1(I)
77	240.5	2.9	1497	2	I48607	procollagen type IV
78	240.5	2.9	1346	1	CGH02E	collagen alpha 1(I)
79	239.5	2.9	920	2	A45748	collagen alpha 1(I)
80	239.5	2.9	1549	2	I48103	type VII collagen
81	239	2.9	1460	1	EDBE1F	immediate-early pr
82	238.5	2.9	627	2	A44112	type VII collagen
83	238.5	2.9	1549	2	I48103	type VII collagen
84	238	2.9	1428	2	T08852	lustrin A - Galito
85	238	2.9	1758	2	T29350	hypothetical prote
86	238	2.9	1759	2	T29351	collagen alpha 2(I)
87	238	2.9	1791	2	T02345	hypothetical prote
88	237.5	2.9	461	2	A31858	ribonuclease-angio
89	237.5	2.9	1344	1	A39175	mucin 1 precursor,
90	237.5	2.9	13288	2	T03099	mucin, submaxillar
91	237	2.9	1446	1	A45344	immediate-early pr
92	236.5	2.9	1763	2	S16366	collagen alpha 2(I)
93	235.5	2.9	1106	2	J00405	hypothetical 119.5
94	235.5	2.9	1106	2	J00405	hypothetical 119.5
95	235.5	2.9	1759	2	T29351	collagen alpha 2(I)
96	235	2.8	1712	1	CGH02B	collagen alpha 1(I)
97	234.5	2.8	931	2	S13580	collagen alpha 1(I)
98	234.5	2.9	1758	2	T29350	hypothetical prote
99	234	2.8	1691	1	CGH06B	collagen alpha 6(I)
100	233.5	2.8	1669	1	CGH04B	collagen alpha 1(I)
101	232	2.8	456	2	S20597	ribonuclease inhib
102	231.5	2.8	1707	2	A35526	collagen alpha 2(I)
103	231	2.8	754	2	A55267	collagen alpha 5(I)
104	231	2.8	920	2	A45748	collagen alpha 1(I)
105	229.5	2.8	1585	2	T31611	hypothetical prote
106	229.5	2.8	3198	2	A44246	collagen alpha 2 F
107	228	2.8	825	2	JC4163	DNA-binding protei
108	228	2.8	1774	2	B56101	collagen alpha 1(I)
109	226.5	2.7	1532	2	A61262	collagen alpha 1(I)
110	226.5	2.7	1707	2	A35526	collagen alpha 2(I)

111	226	2.7	1315	2	A56101	collagen alpha 1(I)
112	226	2.7	1360	2	T39522	hypothetical prote
113	226	2.7	1024	2	S18251	collagen alpha 1(I)
114	223.5	2.7	1027	2	S28774	collagen alpha cha
115	223	2.7	1024	2	S18251	collagen alpha 1(I)
116	222.5	2.7	1446	1	A45344	immediate-early pr
117	222	2.7	671	1	CGRT1S	collagen alpha 1(I)
118	222	2.7	1574	2	T13954	MEGF6 protein - ra
119	222	2.7	1752	2	A43407	collagen alpha 3(I)
120	221.5	2.7	535	2	T48102	RAN GTPase activat
121	221.5	2.7	964	1	CGH02S	collagen alpha 2(I)
122	221.5	2.7	2142	2	B35098	MHC class III hist
123	221	2.7	506	2	A45841	T-complex-associat
124	221	2.7	789	2	T52067	hypothetical prote
125	218.5	2.6	2232	2	T34434	hypothetical prote
126	218	2.6	677	2	S23286	collagen alpha 1(I)
127	218	2.6	744	2	S13435	collagen alpha 1(I)
128	218	2.7	1712	1	CGH02B	collagen alpha 2(I)
129	217	2.6	533	2	T52063	ran GTPase-activat
130	217	2.6	1870	2	S37671	MHC class III hist
131	216.5	2.6	1872	2	S36152	MHC class III hist
132	216.5	2.6	2321	2	S78549	notch3 protein - h
133	215.5	2.6	743	1	S23779	collagen alpha 1(I)
134	215.5	2.6	921	2	S42617	collagen alpha 1(I)
135	215	2.6	635	2	A57131	collagen alpha 2(I)
136	214.5	2.6	635	2	A57131	collagen alpha 2(I)
137	214.5	2.6	1344	1	A55175	collagen alpha 2(I)
138	214	2.6	1184	2	S50832	mucin 1 precursor,
139	213.5	2.6	920	2	B34493	collagen alpha 1(I)
140	213.5	2.6	931	2	S13580	collagen alpha 1(I)
141	213	2.6	1075	2	T31668	hypothetical prote
142	213	2.6	1460	1	EDBE1F	immediate-early pr
143	212.5	2.6	631	2	CG9243	protein F28C1.3 [I
144	212.5	2.6	631	2	T21471	hypothetical prote
145	212.5	2.6	1574	2	T13954	MEGF6 protein - ra
146	212.5	2.6	2796	2	JC4743	fatty-acid synthas
147	212	2.6	825	1	EDBE1F	immediate-early pr
148	211.3	2.6	1964	2	T09059	notch4 - mouse
149	211	2.6	730	2	A36226	collagen alpha 1 c
150	211	2.6	3570	2	T45025	mucin MUC5B, trach

collagen alpha 1(I)  
hypothetical prote  
collagen alpha 1(I)  
collagen alpha cha  
collagen alpha 1(I)  
immediate-early pr  
collagen alpha 1(I)  
MEGF6 protein - ra  
collagen alpha 3(I)  
RAN GTPase activat  
collagen alpha 2(I)  
MHC class III hist  
T-complex-associat  
hypothetical prote  
hypothetical prote  
collagen alpha 1(I)  
collagen alpha 1(I)  
collagen alpha 2(I)  
ran GTPase-activat  
MHC class III hist  
MHC class III hist  
notch3 protein - h  
collagen alpha 1(I)  
collagen alpha 1(I)  
collagen alpha 2(I)  
collagen alpha 2(I)  
mucin 1 precursor,  
atropin-1 - human  
collagen alpha 1(I)  
collagen alpha 1(I)  
hypothetical prote  
immediate-early pr  
protein F28C1.3 [I  
hypothetical prote  
MEGF6 protein - ra  
fatty-acid synthas  
immediate-early pr  
notch4 - mouse  
collagen alpha 1 c  
mucin MUC5B, trach

Search completed: January 7, 2004, 19:47:12  
Job time : 478.493 secs

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 7, 2004, 19:24:01 ; Search time 34,7461 Seconds  
(without alignments)  
12140.333 Million cell updates/sec

Title: US-10-014-269-1  
Perfect score: 8270  
Sequence: 1 gtagacagatccagcgcac.....ataactgttgagtcacaac 4485

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=x1p  
-Q=/cgn2\_1/USPTO\_spool\_p/US10014269/runat\_07012004\_175229\_24504/app\_query.fasta\_1  
.3358  
-DB=SwissProt 41 -OPMT=fastan -SUFFIX=exp -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0  
-DB=bits-Start=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdf -LIST=150  
-DOCLIGN=200 -THR\_SCORE=pc -THR\_PAX=100 -THR\_MIN=0 -ALIGN=90 -MODE=LOCAL  
-OUTFMT=pc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10014269 -GCN1\_1\_97 -runat\_07012004\_175229\_24504 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt 41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Length DB ID Description

1	5448	65.9	1040	1	CARF_HUMAN	Q9hc29	homo sapien
2	4196	50.7	1020	1	CARF_MOUSE	Q9K320	mus musculus
3	1222.5	14.8	953	1	CAR4_MOUSE	Q8bhd0	mus musculus
4	1209.5	14.6	953	1	CAR4_HUMAN	Q9Y239	homo sapien
5	532.5	6.4	1062	1	PYA7_HUMAN	P59046	homo sapien
6	505.5	6.1	1033	1	CIS1_MOUSE	Q8r4b8	mus musculus
7	497	6.0	1034	1	CIS1_HUMAN	Q96p20	homo sapien
8	433	5.2	1130	1	C27A_HUMAN	P33076	homo sapien
9	425	5.1	1155	1	C27A_MOUSE	P79621	mus musculus
10	410.5	5.0	1473	1	NAL4_HUMAN	Q9c000	homo sapien
11	396	4.8	994	1	NAL4_HUMAN	Q96m22	homo sapien
12	385.5	4.7	854	1	PYA5_RAT	Q63035	rattus norvegicus
13	385	4.7	980	1	PYA5_HUMAN	Q9w002	homo sapien
14	379.5	4.6	1062	1	NAL2_HUMAN	P59044	homo sapien
15	366	4.4	892	1	PYA5_HUMAN	Q9tims	mus musculus
16	365.5	4.4	1111	1	NAL2_MOUSE	Q91ws2	mus musculus
17	360.5	4.4	843	1	PYA5_MOUSE	P59047	homo sapien
18	355.5	4.3	1200	1	NAL3_HUMAN	P02461	homo sapien
19	337	4.1	1466	1	CAL3_HUMAN	P08121	mus musculus
20	322	3.9	1464	1	CAL3_MOUSE	P11087	mus musculus
21	320	3.9	1453	1	CAL3_MOUSE	Q01149	mus musculus
22	314.5	3.8	1372	1	CAL2_MOUSE	P02452	homo sapien
23	313	3.8	1464	1	CAL1_HUMAN	P02452	homo sapien
24	312	3.8	1464	1	CAL1_HUMAN	P14728	xenopus laevis
25	310	3.7	1033	1	YHLL_EBV	P59045	homo sapien
26	308.5	3.8	660	1	YHLL_EBV	P03181	epstein-barr virus
27	308	3.7	2944	1	CAL7_HUMAN	Q02388	homo sapien
28	307.5	3.7	1464	1	CAL3_MOUSE	P08121	mus musculus
29	305	3.7	1460	1	CAL1_MOUSE	Q9K477	canis familiaris
30	303	3.7	1466	1	CAL3_HUMAN	P02461	homo sapien
31	301	3.7	1496	1	CAL3_HUMAN	P02458	homo sapien
32	297.5	3.7	1049	1	CAL3_BOVIN	P29400	homo sapien
33	295.5	3.6	1685	1	CAL5_HUMAN	P04258	homo sapien
34	294.5	3.6	1049	1	CAL3_BOVIN	P02453	homo sapien
35	294	3.6	779	1	CAL1_BOVIN	P02458	homo sapien
36	291.5	3.5	1418	1	CAL2_HUMAN	P08123	homo sapien
37	290.5	3.5	1262	1	CAL2_HUMAN	P12105	gallus gallus
38	290	3.5	1690	1	CAL4_HUMAN	P53420	homo sapien
39	289.5	3.5	1364	1	CAL2_BOVIN	P02465	homo sapien
40	288.5	3.5	1364	1	CAL2_BOVIN	P12105	gallus gallus
41	287	3.5	1262	1	CAL3_CHICK	P02466	rattus norvegicus
42	284.5	3.5	1372	1	CAL2_RAT	Q01955	homo sapien
43	282.5	3.4	1670	1	CAL3_HUMAN	Q02388	homo sapien
44	282	3.4	2944	1	CAL1_HUMAN	P02452	homo sapien
45	281	3.4	671	1	CAL1_RAT	P02453	rattus norvegicus
46	281	3.4	1690	1	CAL4_HUMAN	P53420	homo sapien
47	279	3.4	1736	1	CAL2_HUMAN	P13942	homo sapien
48	276	3.4	1453	1	CAL1_CHICK	P02457	gallus gallus
49	276	3.3	1453	1	CAL1_MOUSE	P11087	mus musculus
50	275.5	3.4	1459	1	CAL2_MOUSE	P28481	mus musculus
51	273	3.3	1460	1	CAL1_MOUSE	Q9K477	canis familiaris
52	271	3.3	456	1	RINI_PIG	P10775	sus scrofa
53	270	3.3	1459	1	CAL2_MOUSE	P28481	mus musculus
54	268.5	3.3	1806	1	CAL2_HUMAN	P12107	homo sapien
55	267	3.3	1395	1	CAL2_RANCA	Q42350	rana catesbeiana
56	266.5	3.2	1736	1	CAL2_HUMAN	Q64733	homo sapien
57	266	3.2	1464	1	CAL1_HUMAN	P02452	homo sapien

58	263.5	3.2	1496	1	CA25_HUMAN	P05997 homo sapien
59	263	3.2	1418	1	CA12_HUMAN	P02456 homo sapien
60	263	3.2	1669	1	CA14_MOUSE	P02463 mus musculus
61	262	3.2	1453	1	CA11_CHICK	P02457 gallus galli
62	258.5	3.1	1366	1	CA21_CANFA	P06392 canis famill
63	258.5	3.1	1763	1	CA24_ASCSU	P27393 ascaris suu
64	258.5	3.1	1838	1	CA5_HUMAN	P09098 homo sapien
65	257.5	3.1	1372	1	CA21_RAT	P02466 rattus norv
66	257	3.1	1779	1	CA11_BOVIN	P02453 bos taurus
67	256.5	3.1	1669	1	CA14_HUMAN	P02462 homo sapien
68	254.5	3.1	1366	1	CA21_CANFA	P06392 canis famill
69	254.5	3.1	1838	1	CA5_HUMAN	P09098 homo sapien
70	253	3.1	1362	1	CA12_CHICK	P02467 gallus galli
71	252	3.0	660	1	YHL1_EBV	P03181 Epstein-bar
72	252	3.1	1603	1	CA1F_HUMAN	P02459 mus musculus
73	251.5	3.0	1364	1	CA21_BOVIN	P02465 bos taurus
74	251.5	3.1	1670	1	CA34_HUMAN	P01955 homo sapien
75	251.5	3.1	1685	1	CA54_HUMAN	P29400 homo sapien
76	250.5	3.0	1804	1	CA1B_MOUSE	P02459 mus musculus
77	250	3.0	747	1	CA12_MOUSE	P01149 mus musculus
78	249.5	3.0	1372	1	CA21_MOUSE	P08123 homo sapien
79	249.5	3.0	1366	1	CA1F_HUMAN	P02459 mus musculus
80	248.5	3.0	1603	1	CA1F_HUMAN	P02459 mus musculus
81	248	3.0	1806	1	CA1B_HUMAN	P02459 mus musculus
82	247	3.0	1362	1	CA21_CHICK	P12107 homo sapien
83	245.5	3.0	1516	1	CA1H_HUMAN	P02467 gallus galli
84	244	3.0	2003	1	CA12_HUMAN	P09060 homo sapien
85	243.5	3.0	747	1	CA12_BOVIN	P02467 gallus galli
86	241.5	2.9	1356	1	CA21_ONCKY	P02459 mus musculus
87	241	2.9	744	1	CA1B_RABIT	P02459 mus musculus
88	240.5	2.9	1804	1	CA1B_MOUSE	P02459 mus musculus
89	239	2.9	1736	1	CA2B_HUMAN	P02459 mus musculus
90	238.5	2.9	627	1	SPD2_NEPCL	P01245 mus musculus
91	238.5	2.9	1255	1	MUC1_HUMAN	P13942 homo sapien
92	238	2.9	1758	1	CA24_CAEEL	P15941 h mucin 1 p
93	237.5	2.9	460	1	RINI_HUMAN	P17140 caenorhabdi
94	237	2.9	1446	1	IE1B_PRTXA	P13489 homo sapien
95	236.5	2.9	921	1	CA19_HUMAN	P20849 pseudorabie
96	236.5	2.9	1763	1	CA24_ASCSU	P20849 homo sapien
97	236	2.9	1355	1	CA21_RANCA	P27393 ascaris suu
98	235.5	2.8	1024	1	CARC_HUMAN	P02350 rana catesb
99	235	2.8	1712	1	CA24_HUMAN	P08572 homo sapien
100	234.5	2.9	1758	1	CA24_CAEEL	P17140 caenorhabdi
101	233.5	2.8	815	1	PYG2_DROME	P03948 drosophila
102	233.5	2.8	1669	1	CA14_MOUSE	P02463 mus musculus
103	232	2.8	436	1	RINI_RAT	P09315 rattus norv
104	232	2.8	1691	1	CA64_HUMAN	P08122 mus sapien
105	231.5	2.8	1707	1	CA24_MOUSE	P08122 mus sapien
106	231	2.8	754	1	CA54_CANFA	P08247 canis famill
107	231	2.8	1003	1	MBD6_HUMAN	P09646 homo sapien
108	230	2.8	1403	1	B1RF_MOUSE	P09646 mus musculus
109	229.5	2.8	1736	1	CA2B_MOUSE	P09646 mus musculus
110	229	2.8	1527	1	CA1H_MOUSE	P09646 mus musculus
111	228	2.8	823	1	5B5_RAT	P03003 rattus norv
112	226.5	2.7	1356	1	CA21_ONCKY	P03484 oncohychnu
113	226.5	2.7	1707	1	CA24_MOUSE	P08122 mus musculus
114	225.5	2.7	1027	1	CAFE_RIEPA	P07054 riftinga pach

c 115	225	2.7	911	1	CA1B_BOVIN	Q28083 bos taurus
c 116	222.5	2.7	1446	1	IE1B_PRTXA	P03479 pseudorabie
c 117	222	2.7	671	1	CA11_RAT	P02454 rattus norv
c 118	221.5	2.7	1516	1	CA1H_HUMAN	P39060 homo sapien
c 119	221.5	2.7	2142	1	BAT2_HUMAN	P48634 homo sapien
c 120	220.5	2.7	1461	1	IE1B_PRTXA	P11675 pseudorabie
c 121	219.5	2.7	743	1	CA1B_MOUSE	P07054 mus musculus
c 122	219.5	2.7	744	1	CA1B_HUMAN	P27058 mus musculus
c 123	218.5	2.6	911	1	CA1B_BOVIN	Q28083 bos taurus
c 124	218.5	2.6	2161	1	SHK1_HUMAN	Q9Y666 homo sapien
c 125	218.5	2.7	2161	1	SHK1_HUMAN	Q9Y666 homo sapien
c 126	218	2.6	1402	1	B1RG_MOUSE	Q9Y666 mus musculus
c 127	218	2.7	1712	1	CA24_HUMAN	P08572 homo sapien
c 128	218	2.6	1758	1	CA14_CAEEL	P17135 caenorhabdi
c 129	216.5	2.6	867	1	SSPO_BOVIN	P98167 bos taurus
c 130	216.5	2.6	2321	1	NTC3_HUMAN	Q9UN47 homo sapien
c 131	215	2.6	635	1	CA28_HUMAN	P25067 homo sapien
c 132	215	2.6	689	1	CA29_HUMAN	Q14055 homo sapien
c 133	214.5	2.6	635	1	TR52_HUMAN	P25067 homo sapien
c 134	214.5	2.6	1324	1	TR52_HUMAN	Q9Y642 homo sapien
c 135	213.5	2.6	921	1	CA19_HUMAN	P20849 homo sapien
c 136	213	2.6	1336	1	W146_HUMAN	Q9C0J8 homo sapien
c 137	212.3	2.6	2167	1	SHK1_RAT	Q9WY48 rattus norv
c 138	212	2.6	825	1	ICP0_HSVZH	P28284 herpes simp
c 139	212	2.6	1255	1	MUC1_HUMAN	P15941 h mucin 1 p
c 140	211.5	2.6	1403	1	B1RA_MOUSE	Q9GK45 mus musculus
c 141	211	2.6	5703	1	MUSE_HUMAN	Q14050 homo sapien
c 142	210	2.6	684	1	CA39_HUMAN	P02462 homo sapien
c 143	210	2.6	1669	1	CA14_HUMAN	Q14050 homo sapien
c 144	208.5	2.5	684	1	CA39_HUMAN	Q05722 mus musculus
c 145	208.5	2.5	921	1	CA18_MOUSE	Q03692 homo sapien
c 146	208	2.5	1660	1	CA1A_HUMAN	P31695 mus musculus
c 147	208	2.5	1964	1	NTC4_MOUSE	Q61001 mus musculus
c 148	208	2.5	3718	1	LM45_MOUSE	Q14031 homo sapien
c 149	207.5	2.5	1691	1	CA64_HUMAN	Q14497 homo sapien
c 150	206.5	2.5	1902	1	SHK1_HUMAN	

# ALIGNMENTS

RESULT 1					
CAFE_HUMAN					
ID	CAFE_HUMAN	STANDARD:	PRT:	1040 AA.	
AC	Q9HCE9; Q96RH5; Q96RH6; Q96RH8;				
DT	28-FEB-2003 (Rel. 41, Created)				
DR	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Caspase recruitment domain protein 15 (Nod2 protein) (Inflammatory				
DE	bowel disease protein 1).				
OS	CARD15 OR NOD2 OR IBD1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), MUTAGENESIS OF LYS-305, AND				
RP	VARIANT ARG-908.				

CC TISSUE=Breast;  
RX PubMed=1087742;  
RA Ogura Y., Inohara N., Benito A., Chen F.F., Yamaoka S., Nunez G.;  
RT "Nod2, a Nod1/Apar-1 family member that is restricted to monocytes and  
activates NF-kappaB.";  
RL J. Biol. Chem. 276:4812-4818(2001).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS CD THR-140; ARG-157;  
RP CYS-235; ARG-248; SER-268; SER-289; ASN-291; VAL-301; TRP-311;  
RP VAL-348; ARG-352; CYS-373; SER-414; LEU-431; VAL-432; LYS-441;  
RP VAL-612; THR-612; TRP-664; TRP-702; CYS-703; CYS-713; GLY-723;  
RP VAL-753; VAL-758; LYS-778; MET-793; LYS-843; SER-853; VAL-863;  
RP ARG-908; ASP-918; ASP-924 AND ILE-955, VARIANTS ULCERATIVE COLITIS  
RP THR-140 AND THR-885, AND VARIANT MET-189.  
RC TISSUE=Leukocyte;  
RX MEDLINE=2129172; PubMed=1385576;  
RA Hugot J.-P., Chamaillard M., Zouali H., Lesage S., Ceard J.-P.,  
RA Belaiche J., Almer S., Tyak C., O'Morain C.A., Gassull M., Binder V.,  
RA Finkel Y., Cortot A., Modigliani R., Laurent-Puig P.,  
RA Gower-Rousseau C., Macry J., Colombel J.-F., Sabaotou M., Thomas G.;  
RT "Association of NOD2 leucine-rich repeat variants with susceptibility  
to Crohn's disease.";  
RT Nature 411:599-603(2001).  
RN [3]  
RP VARIANTS BS GLN-334; TRP-334 AND PHE-469.  
RX MEDLINE=2141844; PubMed=11528384;  
RA McCell-Richard C., Lesage S., Rybojad M., Prieur A.M.,  
RA Manouvrier-Hanu S., Hafner R., Chamaillard M., Zouali H., Thomas G.,  
RA Hugot J.-P.;  
RT "CARD15 mutations in Blau syndrome.";  
RT Nat. Genet. 29:19-20(2001).  
CC -1- FUNCTION: Induces NF-kappaB via RICK (CARD14, RIP2) and IKK-  
gamma. Confers responsiveness to intracellular bacterial  
lipopolysaccharides (LPS).  
CC -1- SUBUNIT: Binds to RICK by CARD-CARD interaction.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative initiation;  
CC Comment=2 isoforms, 1/Nod2 (shown here) and 2/Nod2b, are  
CC produced by alternative initiation. Both isoforms can activate  
CC NF-kappaB. Isoform 1 is more abundant.  
CC -1- TISSUE SPECIFICITY: Monocytes-specific.  
CC -1- DISEASE: Defects in CARD15 are the cause of Blau syndrome (BS)  
CC [MIM:186580], a rare autosomal dominant disorder characterized by  
CC early-onset granulomatous arthritis, uveitis and skin rash.  
CC Crohn's disease (CD) [MIM:266600], a form of remitting  
CC inflammatory bowel disease. CD may involve any part of the  
CC gastrointestinal tract, but most frequently the terminal ileum and  
CC colon. Bowel inflammation is transmural and discontinuous. Crohn's  
CC disease is commonly classified as autoimmune disease.  
CC -1- DISEASE: Defects in CARD15 are a cause of susceptibility to  
CC ulcerative colitis [MIM:191390], a chronic inflammatory bowel  
CC disease. In ulcerative colitis, the inflammation is continuous and  
CC limited to rectal and colonic mucosal layers. Ulcerative colitis  
CC is commonly classified as autoimmune disease.  
CC -1- SIMILARITY: Contains 2 CARD domains.

CC -1- SIMILARITY: Contains 1 NACHT domain.  
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; AF178930; AAC33677.1; -;  
CC EMBL; AF385089; AAK70867.1; -;  
CC EMBL; AF385090; AAK70868.1; -;  
CC EMBL; AJ103140; CAC42117.1; -;  
CC Genem; HNCIC5331; CARD15.  
CC MIM; 603956; -;  
CC MIM; 186580; -;  
CC MIM; 266600; -;  
CC MIM; 191390; -;  
CC InterPro; IPR001315; CARD.  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR007091; LRR\_NIMh.  
CC InterPro; IPR007111; NACHT\_NTPase.  
CC Pfam; PF00560; LRR; 2.  
CC SMART; SM00114; CARD; 1.  
CC PROSITE; PSS0209; CARD; 1.  
CC PROSITE; PSS0837; NACHT; 1.  
CC KW ATP-binding; Repeat; leucine-rich repeat; Disease mutation;  
CC Alternative initiation; Polymorphism.  
CC CHAIN 1 1040  
CC FT CASPASE RECRUITMENT DOMAIN PROTEIN 15,  
CC ISOFORM 1.  
CC FT CHAIN 28 1040  
CC FT INIT MET 28 28  
CC FT DOMAIN 26 122  
CC FT DOMAIN 126 218  
CC FT DOMAIN 293 618  
CC FT NP\_BIND 299 306  
CC FT REPEAT 786 812  
CC FT REPEAT 814 837  
CC FT REPEAT 842 865  
CC FT REPEAT 926 949  
CC FT REPEAT 954 977  
CC FT REPEAT 982 1005  
CC FT VARIANT 140 140  
CC FT VARIANT 157 157  
CC FT VARIANT 189 189  
CC FT VARIANT 235 235  
CC FT VARIANT 248 248  
CC FT VARIANT 266 266  
CC FT /FTid=VAR\_012667.  
CC FT /FTid=VAR\_012668.  
CC FT /FTid=VAR\_012669.  
CC FT /FTid=VAR\_012670.



FT	VARIANT	289	289	N -> S.
FT	VARIANT	291	291	/FTid=VAR_012671.
FT	VARIANT	291	291	D -> N (in CD).
FT	VARIANT	294	294	/FTid=VAR_012672.
FT	VARIANT	294	294	T -> S (in CD).
FT	VARIANT	301	301	/FTid=VAR_012673.
FT	VARIANT	301	301	A -> V (in CD).
FT	VARIANT	311	311	/FTid=VAR_012674.
FT	VARIANT	311	311	R -> W (in CD and ulcerative colitis).
FT	VARIANT	334	334	/FTid=VAR_012675.
FT	VARIANT	334	334	R -> Q (in BS).
FT	VARIANT	334	334	/FTid=VAR_012676.
FT	VARIANT	334	334	R -> W (in BS).
FT	VARIANT	348	348	/FTid=VAR_012677.
FT	VARIANT	348	348	L -> V (in CD).
FT	VARIANT	352	352	/FTid=VAR_012678.
FT	VARIANT	352	352	H -> R (in CD).
FT	VARIANT	373	373	/FTid=VAR_012679.
FT	VARIANT	373	373	R -> C (in CD).
FT	VARIANT	414	414	/FTid=VAR_012680.
FT	VARIANT	414	414	N -> S (in CD).
FT	VARIANT	431	431	/FTid=VAR_012681.
FT	VARIANT	431	431	S -> L (in CD).
FT	VARIANT	432	432	/FTid=VAR_012682.
FT	VARIANT	432	432	A -> V (in CD).
FT	VARIANT	441	441	/FTid=VAR_012683.
FT	VARIANT	441	441	E -> K (in CD).
FT	VARIANT	469	469	/FTid=VAR_012684.
FT	VARIANT	469	469	L -> F (in BS).
FT	VARIANT	612	612	/FTid=VAR_012685.
FT	VARIANT	612	612	A -> T (in CD).
FT	VARIANT	612	612	/FTid=VAR_012686.
FT	VARIANT	684	684	A -> V (in CD).
FT	VARIANT	684	684	/FTid=VAR_012687.
FT	VARIANT	702	702	R -> W (in CD).
FT	VARIANT	702	702	/FTid=VAR_012688.
FT	VARIANT	703	703	R -> W (risk factor for CD).
FT	VARIANT	703	703	/FTid=VAR_012689.
FT	VARIANT	713	713	R -> C (in CD and ulcerative colitis).
FT	VARIANT	713	713	/FTid=VAR_012690.
FT	VARIANT	725	725	R -> C (in CD).
FT	VARIANT	725	725	/FTid=VAR_012691.
FT	VARIANT	755	755	A -> G (in CD).
FT	VARIANT	755	755	/FTid=VAR_012692.
FT	VARIANT	758	758	A -> V (in CD and ulcerative colitis).
FT	VARIANT	758	758	/FTid=VAR_012693.
FT	VARIANT	778	778	A -> V (in CD).
FT	VARIANT	778	778	E -> K (in CD).
FT	VARIANT	793	793	/FTid=VAR_012694.
FT	VARIANT	793	793	E -> K (in CD).
FT	VARIANT	843	843	/FTid=VAR_012695.
FT	VARIANT	843	843	V -> M (in CD).
FT	VARIANT	853	853	/FTid=VAR_012696.
FT	VARIANT	853	853	E -> K (in CD).
FT	VARIANT	863	863	/FTid=VAR_012697.
FT	VARIANT	863	863	N -> S (in CD).
FT	VARIANT	863	863	/FTid=VAR_012698.
FT	VARIANT	863	863	M -> V (in CD).

FT	VARIANT	885	885	/FTid=VAR_012699.
FT	VARIANT	885	885	A -> T (in ulcerative colitis).
FT	VARIANT	908	908	/FTid=VAR_012700.
FT	VARIANT	908	908	G -> R (in CD).
FT	VARIANT	918	918	/FTid=VAR_012701.
FT	VARIANT	918	918	A -> D (risk factor for CD).
FT	VARIANT	924	924	/FTid=VAR_012702.
FT	VARIANT	924	924	G -> D (in CD).
FT	VARIANT	955	955	/FTid=VAR_012703.
FT	VARIANT	955	955	V -> I.
FT	VARIANT	955	955	/FTid=VAR_012704.
FT	VARIANT	955	955	K -> R: NO ACTIVATION.
FT	VARIANT	955	955	SQ: SEQUENCE 1040 AA; 115282 MW; 0037592D96D7DBDF CRC64;

Alignment Scores:

Pred. No.:	2,85e-267	Length:	1040
Score:	5448.00	Matches:	1040
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.88%	Indels:	0
DB:	1	Gaps:	0

US-10-014-269-1 (1-4485) x CARF\_HUMAN (1-1040)

QY	106	ATGAGGGAAGAGGCTGCTACGCTCTACGATGAGGAAAGACGATGCTCTC	165
DB	1	MetGlyGluGluGlySerAlaSerHisAspGluGluArgAlaSerValLeuLeu	20
QY	166	GGACATCTCCGGGTTGTGAAATGCTCGAGAGGCTTTTACGACAGAGGACG	225
DB	21	GlyHisSerProGlyGlyGluMetCysSerGlnAlaPheGlnAlaArgSerGln	40
QY	226	CTGTCAGAGCTGCTGCTCTCAGAGGCTCCGAGAGGCTTCGAGATGTCGAGTGGCTG	285
DB	41	LeuValGluLeuLeuValSerGlySerLeuLeuGlyPheGluSerValLeuAspTyrLeu	60
QY	286	CTGTCTGGAGGCTCTCTCTCGAGAGCTACGAGGCTTCACCTCTGAGGACGCT	345
DB	61	LeuSerTyrGluValLeuSerTyrGluAspTyrGluGlyPheHisLeuLeuGlyGlnPro	80
QY	346	CTCTCCGACTGGCCAGGCGCTTCGAGACCGCTCGAATTAAGGGTACTTGGGCTGT	405
DB	81	LeuSerHisLeuAlaArgAlaGluLeuAspThrValTyrAsnGlyGlyTyrTyrAlaCys	100
QY	406	CAGAGCTCATCGCGCTGCCAAGAACCCAGGCCGACACGCAATGCCCCAGCTGCAT	465
DB	101	GlnLysLeuIleAlaAlaGlnGluAlaGlnAlaAspSerGlnSerProLysLeuHis	120
QY	466	GGCTGCTGGACCCCACTGCTCCACCCAGCCGAGCTCGAGNGTCACCGGCGCAGGC	525
DB	121	GlyCysTyrAspProHisSerLeuHisProLysArgAspLeuGlnSerHisArgProAla	140
QY	526	ATTGTACAGAGGCTCCACAGCAATGTGAGGACATCTGACCTGGCATGGAGCGGAGT	585
DB	141	IleValArgTyrGluHisSerHisValGluLysMetLeuAspLeuAlaTyrGluArgGly	160
QY	586	TTCGTACACCAAGTATGAAATGTGATGAATTCAGGTTCGAGATCTTCACACCGTCCACAG	645

Db	161	Phenylserine	180
QY	646	GAAAGAGCTGCTTGAATCTTCCACCGGTAAAGCATGATGGCTGCTTCTTCA	705
Db	181	AlaArgLeuLeuAspLeuAlaThrValIysAlaAsnGlyLeuAlaAlaPheLeuLeu	200
QY	706	CAACATGTCAGGAATACCACTCCCACTGGCCCTGGCTTGGAAAGCTGCACATGCAAG	765
Db	201	GlnHisValGlnGlnLeuProValProlLeuAlaLeuProlLeuGlnAlaAlaThrCysIys	220
QY	766	AAGTATGGCCAGCTGAGGACCAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	825
Db	221	LysThrMetAlaIysLeuArgThrThrValSerAlaGlnSerArgPheLeuSerThrTyr	240
QY	826	GATGAGCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	885
Db	241	AspGlyAlaGlnThrLeuCysLeuGlnAspIleTyrThrGlnAsnValLeuGlnValTyr	260
QY	886	GCAAGATGTGGGATGGCTGGAACCCCGCAGAAAGCCACAGCCCTGGGCTGGAGAG	945
Db	261	AlaAspValGlnMetAlaGlyProGlnIysSerProAlaThrLeuGlnLeuGln	280
QY	946	CTCTGACGACCCCTGGCCACTCATAGACATGCGGACATGCTGCTGGGTGGTGGAG	1005
Db	281	LeuPheSerThrProGlyHISLeuAsnAspAlaAspThrValLeuValValGlyGln	300
QY	1006	GCGGCGAGTGGCAGAGACAGCTCTGACGCGCTGCACTGCTGTGGCTGCGGCGAA	1065
Db	301	AlaGlySerGlyLysSerThrLeuLeuGlnAlaGlnHisLeuLeuThrAlaGlyGln	320
QY	1066	GACTCCAGGAATTTCTTTTGTCTTCCCATTCAGCTGCCGGACGCTGCAGTGCATGGCC	1125
Db	321	AspPheGlnGlnPheLeuPheValPheProPheSerCysArgGlnLeuGlnCysMetAla	340
QY	1126	AAACCACTCTCTGGCGGACTCTACTCTTTAGACATGCTGTGGCTGATGTTGATGAC	1185
Db	341	LysProlLeuSerValArgThrLeuLeuPheGlnHisCysTyrProAspValGlyGln	360
QY	1186	GAAAGACATCTTCCAGTTACTCTTGAACACCTGACCGTGTCTGTTAACTTTGATGGC	1245
Db	361	GlnAspIlePheGlnLeuLeuLeuAspHisProAspArgValLeuLeuThrPheAspGly	380
QY	1246	TTTGACGAGTTCAAGTTCAGGTTCAAGTGTGTAAGCCACTGCTCCCGACGACGCC	1305
Db	381	PheAspGlnPheLysPheArgPheThrAspArgGlnArgHisCysSerProThrAspPro	400
QY	1306	ACCTGTGTCAGACCGCTGCTTCAACTTTCGACGGGCAACTGCTGAAGATGCCCGC	1365
Db	401	ThrSerValGlnThrLeuLeuPheLeuLeuGlnGlnIysAsnLeuLeuLysAsnAlaArg	420
QY	1366	AAAGTGTGACCAAGCGTCCGCGCGCTGTGTGGCGTTCCTCAAGAAAGATCCGACCC	1425
Db	421	LysValValThrSerArgProAlaAlaValSerAlaPheLeuArgGlyTyrIleArgThr	440
QY	1426	GAGTTCACCTCAAGGGCTTCTGTAACGAGGCACTGACGTGCTGAGGAAGCGCAAT	1485
Db	441	GlnPheAsnLeuLysGlyPheSerGlnGlnGlnIylGlnLeuTyrLeuArgLysArgHis	460
QY	1486	CATGACCCCGGGTGGCGGACCGCTCATCCGCGCTGCTCAAGACATCAAGCCCTGCAC	1545
Db	461	HisGlnProGlyValAlaAspArgLeuIleArgLeuLeuGlnGlnThrSerAlaLeuHis	480
QY	1546	GGTTGTGCAACTGCTGCTTCTCATGATGGTGTCCAAATGCCACCAAGAACTGCTG	1605
Db	481	GlyLeuCysHisLeuProValPheSerThrPheValSerLysCysHisSerGlnLeuLeu	500
QY	1606	CTGACAGAGGGGGGGTCCCAAAAGCACTACAGATATGTACCTGCTGATTCGACGAT	1665
Db	501	LeuGlnGlnGlyLysSerProLysThrThrThrAspMetTyrLeuLeuLeuGlnHis	520
QY	1666	TTTCTGCTGCATGCAACCCCGCCAGACTCACTTCCCAAGGTCTGGAGCCAGTCTTCTT	1725
Db	521	PheLeuLeuHisAlaThrProProAspSerAlaSerGlnGlyLeuGlyProSerLeuLeu	540
QY	1726	CGGGGCGCGCTCCCAACCTCTGCACTGGGCAAGTGGCTCTGTGGGGCTGGCATG	1785
Db	541	ArgGlyArgLeuProThrLeuLeuHisLeuGlnArgLeuAlaLeuThrPheGlyMet	560
QY	1786	TGCTGCTACGTGTCTCAAGCCAGACCTCCAGGCAAGCTCAAGCTGATGACATT	1845
Db	561	CysCysTyrValPheSerAlaGlnGlnLeuGlnAlaAlaGlnValSerProAspAspIle	580
QY	1846	TCTCTTGGCTCTGGTGGTGGCAAGGTGTGTGCGGAGGATGAGGCGCCCGCTGAA	1905
Db	581	SerLeuGlnPheLeuValArgAlaLysGlyValValProGlySerThrAlaProlLeuGln	600
QY	1906	TTCCTTCAATCACTTTCAGGTGCTTTTCCCGGCTTCTACCTGGCACTCAAGTCAAT	1965
Db	601	PheLeuHisIleThrPheGlnCysPhePheAlaAlaPheTyrLeuAlaLeuSerAlaAsp	620
QY	1966	GTCCACCAAGTTGCTGACACACCTTCAATTGGGAGGCGCACTCAACCAAG	2025
Db	621	ValProProAlaLeuLeuArgHisLeuPheAsnCysGlyArgProGlyLysAsnSerProMet	640
QY	2026	GCCAGGCTCTGCCCAAGATGTGATCCAGGCGCTCGAGAGGAAAGACACAGAGGTGGCA	2085
Db	641	AlaArgLeuLeuProThrMetCysIleGlnAlaSerGlnGlyLysAspSerSerValAla	660
QY	2086	GCTTGTGCGAGAAGGCCAGCCGACCAACCTTCAGATCAACAGACGCTTCTGCGAGG	2145
Db	661	AlaLeuLeuGlnLysAlaGlnProHisAsnLeuGlnIleThrAlaAlaPheLeuAlaGly	680
QY	2146	CTGTGTCCCGGAGACACTGGGGCGCTGCTGAGTGGCAGACATCGAAGAGCGCGT	2205
Db	681	LeuLeuSerArgGlnHisIleTyrGlyLeuLeuAlaGlnCysGlnThrSerGlnLysAlaLeu	700
QY	2206	CTCCGGCGCAAGGCTGTGCCCGCTGCTGTGCTGAGCCCGACGCTCCGCAAGCACTTCAC	2265
Db	701	LeuArgArgGlnIleCysAlaArgTyrCysLeuAlaArgSerLeuArgLysHisPheHis	720
QY	2266	TCCATCCCGGCACTGCAACCGGGTGGGCGAAGGGTGCATGCGTGGCGGGTTCATC	2325
Db	721	SerIleProProAlaIleProGlyGlnAlaLysSerValHisAlaIleProGlyMetIle	740

QY 2326 TGGCTCATCOGAGGCTGTACAGAGATGCAAGAGAGGCGCTGCGAAGGCTGCACGT 2385  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 741 TrrpleuilearfgserleuTyrglwmecglnluglnrghleuylaaqlyslalaalarg 760  
 QY 2386 GGCCCTGATGTTGGGACACCTCAAGTTGACATTTTGCAAGTGGGGCCCACTGAGTGTCT 2445  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 761 GlyleuamValglYhstleuylleuTrrhlecysSerValglCyrprothrlnucylala 760  
 QY 2446 GCCCTGGCTTTGTGCTGACAGACCTCCGGCGCCCTGGCCCTGCAAGCTGACACAC 2505  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 781 AlaleuAlapheValleuGlnhstleuArgprValAlaleuGlnleuAspTyrasn 800  
 QY 2506 TCTGGGTGACATTGGCGTGGAGACACTGCTGCCCTTGGCTGGTGTCTGACAGGCTCTG 2565  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 801 SerValglYAspIleaglYalGlnInleuLeuprocysleuGlYalCysYslAlleu 820  
 QY 2566 TATTGGCCATTAACAAATCTCAAGACCAAGGCACTTCAAGACTCATTAATGTCTCTT 2625  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 821 TyrleuArgAspAsnAsnIleSerAspArgGlyIleCysYslleuIleGlnCysAlleu 840  
 QY 2626 CACTCGAGCAATTGCAAGATTAGCTTATTAACAAACAATTGACTGAGCGCTGGCA 2685  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 841 HisCysGlnGlnInleuGlnhstleuAlleuPheAsnAsnIlyleuThrAspGlyCysAla 860  
 QY 2686 CACTCCATGGCTAAAGCTCCTTGCATGCAAGGCAAACTTCTTGGCATTGAAGCTTGGGAAT 2745  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 861 HisSerMetAlaIlyleuAlleuAlaCysArgGlnAsnPheleuAlleuArgleuGlyAsn 880  
 QY 2746 AACTCATGCTGCGCGGAGGCCAAGTGTGGCGCAAGGGGCTCCAGAGCAACCTCTCC 2805  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 881 AsnTyrIleThrAlaAlaGlyAlaGlnValleuAlaGlnGlyleuArgGlyAsnThrSer 900  
 QY 2806 TTGCAGTTCTGGGATTCTGGGGCCACAGAGTGGGTGAGAGGGGCCCGCCCTGCT 2865  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 901 LeuGlnPheleuGlyPheTrpGlyAsnArgValGlyYAspGlnGlyAlaGlnAlaAlleuAla 920  
 QY 2866 GAAAGCTTGGGTGATCAACAAGCTTGAAGTGGCTCAAGCTTGGGGCAACAATTGGC 2925  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 921 GlnAlaLeuGlyAspHisGlnSerleuArgTrpIleuSerleuValGlyAsnAsnIleGly 940  
 QY 2926 AGTGTGGTGGCCAAAGCTTGGCACTGATGCTGGCAAGAAGCATGCTGTAGAGAACTC 2985  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 941 SerValglYAlaGlnAlaAlleuAlaIleuMetleuAlaIlyAsnValIleleuGlnGlyleu 960  
 QY 2986 TGGCTGAGAGAAACAATCTCAAGATGAAGGTATGTTCTTCGCAAGAGACTGAG 3045  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 961 CysleuGlnGlnAsnHisleuGlnAspGlnGlyValCysSerleuAlaGlnGlyleuYls 980  
 QY 3046 AAAAATCAAGTTGAAAATCCTGAAGTTGTCAATTAAGTCACTCACTCACTGAGGGA 3105  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 981 LysAsnSerSerleuYlsIleleuYlsleuSerAsnAsnGlyIleThrTyrleuGlyAla 1000  
 QY 3106 GAAAGCTTCTGCAAGGCTTGAAGAGATGAACAATCTGAGAGTCTGGCTCGAGGG 3165  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1001 GlnAlaLeuGlnAlaIleuGlnArgAsnAspThrIleleuGlnValIleuAlaGly 1020

QY 3166 AACACTTCTCTAAGAGAGATTGACAAAGCTCGAGGAGCAACAAGCTTGTCT 3225  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1021 AsnThrPheSerleuGlnGlnValAspIlyleuGlyCysArgAspThrArgleuLeu 1040  
 RESULT 2  
 ID CARF\_MOUSE STANDARD; PRT; 1020 AA.  
 AC Q8K3ZF.  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Caspase recruitment domain protein 15 (Nod2 protein).  
 GN CARD15 OR NOD2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 (1)  
 SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=BALB/c; TISSUE=Monocytes;  
 RA Itanaga Y., Davey M.P., Martin T.M., Planck S.R., DePirast M.L.,  
 RA Baugh M.M., Suing C., Rosenbaum J.T.;  
 RT "Cloning, sequencing and expression analysis of the murine Nod2/CARD15  
 RT gene";  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 (2)  
 SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANTS ALA-212; ARG-240;  
 RP CYS-422; VAL-485; ALA-603; ILE-675 AND GLN-925.  
 RC STRAIN=NRR1; TISSUE=Breast cancer;  
 RA MEDLINE=22389297; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.R., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.U., Ueda T.B., Tomshyuk S., Carninci P., Prange C.,  
 RA Rhee S.S., Loughlano N.A., Peters G.J., Abrahams R.D., Mullany S.J.,  
 RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Holyc S.W.,  
 RA Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Induces NF-kappaB via RICK (CARD14, RIP2) and IKK-  
 CC gamma. Confers responsiveness to intracellular bacterial  
 CC lipopolysaccharides (LPS) (By similarity).  
 CC -1- SUBUNIT: Binds to RICK by CARD-CARD interaction (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;

[illegible]

Db	246	AlaGlyAlaLeuGlnIysSerProAlaIleLeuGlyLeuGlnAlaSerLeuPheLeuPheAspThrHis	265
QY	961	GGCCAACTCAATGACAAATGCGGAACGTGGCTGGTGGGTGAAGCGGGGACGTGGCAAG	1020
Db	266	GLYHISLeuAsnArgAspAlaAspThrIleLeuValGlyGlyAlaGlySerGlyIys	285
QY	1021	AGCAAGCTCTGCAGCGGAGTGCACCTGCTGGTGGCTGCGAGGCAAGACTTCGAGAAATT	1080
Db	286	SerThrLeuLeuGlnArgLeuHISLeuLeuThrPAlaThrGlyArgSerPheGlnGluPhe	305
QY	1081	CTCTTTGCTCCCATTCAGCTGCGCGGACGCTGCAGTGCAGGCGAAACAACTCTCGTG	1140
Db	306	LeuPheIlePheProPheSerCysArgGlnLeuGlnCysValAlaIlyIysProLeuSerLeu	325
QY	1141	CGAAGTACTCTTTAGACACTGCTGTTGGCTGATGTTGGTCAAGAAAGACATTTCCAG	1200
Db	326	ArgThrLeuLeuPheGlnHISCysCysTrpProAspValAlaGlnAspAspValPheGln	345
QY	1201	TTCACCTTGAACCAACCTGACCGGTGCTGTTAAACCTTGATGGCTTTGAGAAATTCAAG	1260
Db	346	PheLeuLeuAspHisProAspArgValAlaLeuLeuThrPheAspGlyLeuAspGlyLeuPheIys	365
QY	1261	TTCAAGTTCACGAGATCGTGAACCGCACTGCTCCCGCAACGACCCCAACTGTGCAGAC	1320
Db	366	PheArgPheThrAspArgGlyAlaArgHISCysSerProIleAspProHisSerValGlnThr	385
QY	1321	CTGCTTTCAACCTCTGCAGGGCAACCTGGCTGAAGAAATGCCGGAAGGTGTGTACACAG	1380
Db	386	LeuLeuPheAsnLeuLeuGlnGlyAsnLeuLeuIlyAsnAlaCysIlyValLeuThrSer	405
QY	1381	CGTCCGCGCGGTGTGTGCGCTCTCTCGAGAGATGACATCCGCAACGATTCACAACTCAAG	1440
Db	406	ArgProAspAlaValSerAlaLeuLeuArgIlyPheValAlaGlnThrGluLeuGlnLeuIys	425
QY	1441	GGCTCTCTGGAACAGGGGATCGAGGTGTAACCTGCAGAGAACCGCATGACGACCGGGGGT	1500
Db	426	GlyPheSerGluGlnGlyIleGlnLeuIlyLeuArgIlyHisHisSerGlyLysProGlyAla	445
QY	1501	CGGAGACCGGCTCATCCGCGTCTCTCGAGAGACTCAAGCCCTGCACGGTTTGTGCACCTG	1560
Db	446	AlaAspArgLeuIleGlnLeuIleGlnAlaThrSerAlaLeuHISGlyLeuCysHisLeu	465
QY	1561	CGTGTCTCTCATGATGATGATGCACAAATGCCACGAAACTGTTGCTGCAGAGGGGGG	1620
Db	466	ProValPheSerTrpPheValSerArgCysHisSerGlyLeuLeuLeuLeuGlnHisAsnGly	485
QY	1621	TCCCAAAAGACACTACAGATATGACTGCTGATCTTCGACGACATTTTCTGCTGACTGCC	1680
Db	486	PheProThrIleSerThrAspMetIlyLeuLeuIleLeuGlnHisPheLeuLeuHISAla	505
QY	1681	AACCCCCCAAGACTCCCAAGAGCTCTGGGACCGCAAGTCTCTCGAGGGCGGCTCCCC	1740
Db	506	SerProProAspSerSerProLeuGlyLeuGlyProGlyLeuLeuIleSerArgLeuSer	525
QY	1741	AACCTCTGCACCTGGGAGAGACTGGCTGTGGGGGCTGGGCATGTGCTGCTACGTGTC	1800
Db	526	ThrLeuLeuHISLeuGlyHisLeuAlaLeuLeuGlyLeuLeuHISMetSerCysTrpValPhe	545
QY	1801	TGAGCCAGACGCTCCAGGACCAACAGCTTCAGCCCTGATGACATTTCTCTGGCTCGTG	1860
Db	546	SerAlaGlnIleLeuGlnAlaAlaGlnValAspSerAspAspIleSerLeuIlyPheLeu	565
QY	1861	GTCGCTGCCAAAGGTGCTGCCAGGAGATAGAGCGCCCTGAAATCTTCACATCACT	1920
Db	566	ValArgAlaGlnSerSerValProGlySerIlyAlaProLeuGlnPheLeuHISIleThr	585
QY	1921	TTCAGTGCCTTTTGCCCGGCTCTACCTGGCACTCAGTGCATGTGCCACAGCTTGG	1980
Db	586	PheGlnCysPhePheAlaIlePheIlyLeuAlaValSerAlaAspThrSerValAlaSer	605
QY	1981	CTCAGACACTTCATATGTGGACGCGGCAACACTCAACAAAGGACGGCTCGTGGC	2040
Db	606	LeuIysHisLeuPheSerCysIlyArgLeuGlySerSerLeuLeuIlyArgLeuLeuPro	625
QY	2041	ACAAATGTCATCCAGGCTCGGAGGAAAGACAGACGGTGGCAGCTTGTCTGACAGAG	2100
Db	626	AsnLeuCysIleGlnIlySerTrpValIlySerGluAlaIleLeuLeuGlnIys	645
QY	2101	GCGAGCGCCCAACCTTCAGATCAAGACAGGCTTCCTGCAGGGCTGTTGTCCGCGAG	2160
Db	646	AlaGluProHisAsnLeuGlnIleThrAlaAlaPheLeuAlaGlyLeuLeuSerGlnGln	665
QY	2161	CACGTGGGCTGCTGGCTGATGGCAACATCTGAAGAAAGCCCTGCTCCGCGCCACGCC	2220
Db	666	HisArgAspLeuLeuAlaAlaCysGlnValSerGlyAlaValLeuLeuGlnArgGlnAla	685
QY	2221	TGTGCCCGTGTGTGTGCGCCGACGCTCCGCAAGGACTTCTCACTTCAATCCCGCAAGT	2280
Db	686	ArgAlaArgSerCysLeuAlaHisSerLeuArgIleHisPheHISerIleProAla	705
QY	2281	GACCCGGGTGAGGCAAGAGGCTGATGCAATGCGCGGGTTCATCTGCTCATCCGACAG	2340
Db	706	ValProGlyGlyThrIlySerPheHisAlaIleProGlyPheIleTrpLeuIleArgSer	725
QY	2341	CTGACAGATGCAGAGAGAGCGGCTGCTCGAAGGCTGCACCTGAGCTGAATGTGGG	2400
Db	726	LeuIlyGluMetGlnGlnGlnLeuAlaGlnGlyAlaValAlaArgValAspIleGly	745
QY	2401	CAACTCAATGACATTTGCAAGTGTGGGCGCCCATGAGTGTGCGCTGCGCTTTGTG	2460
Db	746	HisLeuIlyLeuThrTrpPheCysArgValGlyProAlaGlyCysAlaAlaLeuAlaPheVal	765
QY	2461	CTGCAGACCTCCGGCGCGCCGTGGCTCGACGCTGAAGTCAACATCTGTGGGTGAACAT	2520
Db	766	LeuGlnHisLeuGlnArgProValAlaLeuGlnLeuAspIlyArgAsnSerValGlyAspVal	785
QY	2521	GGCTGAGACGACTGCTGGCTTGGCTGCTGCTGCTGCAGAGGCTCGAATTTGCGCAATAC	2580
Db	786	GlyAlaGlnGlnLeuArgProCysGlnGlyAlaCysThrAlaLeuIlyTrpLeuArgAspAsn	805
QY	2581	AATATCTACACCGAGGACCTGCAGACTCATGAATGTGCTCTCACTGGACGACATTTG	2640
Db	806	AsnIleSerAspArgGlyAlaArgThrLeuValGlnCysAlaLeuArgCysGlnGlnLeu	825

QY 2641 CAGAGTTAGCTTATTCACAAATGACTGACGCTGTCACACTCCATGGCTAAG 2700  
 Db 826 GlnysLeuAlaLeuPheAsnAsnLysLeuThrPalaCysAlaCysSerMetAlaLys 845  
 QY 2701 CTCTTCATGACGAGAACTCTTGCACTTGGAGGCTGGGGAATAACTCATCTACAGGCC 2760  
 Db 846 LeuLeuAlaHisLysGlnAsnPheLeuSerLeuArgAlaGlnAsnHisLeuAla 865  
 QY 2761 GCGGAGACCCAAAGTGTGCGCAAGGCTCCAGGCAACACTTGGAGTTCCTGGGA 2820  
 Db 866 AlaGlyAlaGluValLeuAlaGlnGlyLeuLysSerAsnThrSerLeuLysPheLeuGly 885  
 QY 2821 TTCTGGGGCAACAGAGTGGGTGACGAGGGGGCCCTGGCTGGAAGCTTGGGTGAT 2880  
 Db 886 PheTrpGlyAsnSerValGlyAspIysGlyThrGlnAlaLeuAlaGluValAlaAsp 905  
 QY 2881 CACCAAGCTTGAAGTGGCTCAGGCTGGTGGGGAACAATTGGAGTGTGGGTCCCA 2940  
 Db 906 HisGlnAsnLeuLysTrpLeuSerLeuValGlyAsnAsnLysGlySerMetGlyAlaGlu 925  
 QY 2941 GCCTGGCACTGATGCTGGCAAGAAAGACGTCAAGAGAACTCTGCTGGAGAGAAAC 3000  
 Db 926 AlaLeuAlaLeuLeuLeuGluLysSerLeuGluLysLeuLysLeuLysLeuLys 945  
 QY 3001 CATCTCCAGATGAAGGTGTATGTTCTCTGCAAGAGACTGAGAAATAATCAAGTTG 3060  
 Db 946 HisLeuCyAspGlnGlyValLysSerLeuAlaGlnGlyLeuLysArgAsnSerThrLeu 965  
 QY 3061 AAAATCCTGAAGTGTCCAAATACGTCACTACCTAGGGGCAAGAGGCTCTGCAAG 3120  
 Db 966 LysPheLeuLysLeuSerAsnAsnGlyLeuThrArgGlyAlaGluAlaLeuLeuGln 985  
 QY 3121 GCCCTTGAAGGAATGACACATCTGCAAGCTGAGCTCCAGGGAACACTTCTCTCTA 3180  
 Db 986 AlaLeuSerArgAsnSerAlaLeuGluValTrpLeuArgGlyAsnThrPheSerLeu 1005  
 QY 3181 GAGGAGGTGACAAAGCTCGGCTGCAAGGACACAGACTCTTCTT 3225  
 Db 1006 GlnGluLeuGlnThrLeuSerSerArgAspAlaArgLeuLeu 1020

Search completed: January 7, 2004, 19:42:15  
 Job time : 424.746 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 7, 2004, 19:25:21 ; Search time 170.731 Seconds  
 (without alignments)  
 13537.774 Million cell updates/sec

Title: US-10-014-269-1  
 Perfect score: 8270  
 Sequence: 1 gtagagcagatccagctccac.....ataaacgttgagtcacaac 4485

Scoring table:

ELOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Command line parameters:  
 -MODEL=frame+\_n2p.model -DEV=xlp  
 -Q/cgn2\_1/USP10.spool\_p/US10014269/runat\_07012004\_175229\_24521/app\_query.fasta\_1  
 .3338  
 -DB=SPTREMBL\_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=150  
 -DOCLIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL  
 -OUTFMT=ptc -NORW=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=200000000  
 -USER=US10014269.GCEN.1.1.482.G-unatc.07012004.175229.24521 -NCPU=6 -ICPU=3  
 -NO\_MAP -LARGEOUTERY -NEG\_SCORES=0 -WALT -DSBLOCK=100 -LONCLDG  
 -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*

11: sp\_rodent:\*\n12: sp\_virus:\*\n13: sp\_vertebrate:\*\n14: sp\_unclassified:\*\n15: sp\_rvifus:\*\n16: sp\_bacteriap:\*\n17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	4186	50.7	1020 11 Q8K320	Q8K320 mus musculus
2	1222.5	14.8	953 11 Q8BHB0	Q8BHB0 mus musculus
3	1209.5	14.6	953 4 Q8BWF5	Q8BWF5 homo sapien
4	1019	12.3	706 11 Q8BUT6	Q8BUT6 mus musculus
5	505.5	6.1	1033 11 Q8R438	Q8R438 mus musculus
6	483	5.8	892 4 Q8NEU4	Q8NEU4 homo sapien
7	437	5.3	1052 7 Q8CUD8	Q8CUD8 rattus norv
8	437	5.3	1073 7 Q8CUD9	Q8CUD9 rattus norv
9	437	5.3	1153 7 Q8GJED0	Q8GJED0 rattus norv
10	425	5.1	1155 7 Q8PPE1	Q8PPE1 mus musculus
11	410.5	5.0	778 4 Q8NFA8	Q8NFA8 homo sapien
12	377.5	4.6	713 6 Q8SIE7	Q8SIE7 macaca fasc
13	365.5	4.4	499 4 Q8REU3	Q8REU3 homo sapien
14	360.5	4.4	982 11 Q8BUD40	Q8BUD40 mus musculus
15	357.5	4.3	733 4 Q8TEE2	Q8TEE2 homo sapien
16	353	4.3	5146 6 Q8SPM4	Q8SPM4 bos taurus
17	350.5	4.2	1056 4 Q8NFA2	Q8NFA2 homo sapien
18	338.5	4.1	5146 6 Q8SPM4	Q8SPM4 bos taurus
19	332	4.0	846 4 Q8IXT0	Q8IXT0 mus musculus
20	329.5	4.0	1787 10 Q8MAX9	Q8MAX9 chlamydomon
21	325	4.0	1464 11 Q8BLM4	Q8BLM4 mus musculus
22	322	3.9	932 4 Q8KIL4	Q8KIL4 homo sapien
23	321.5	3.9	660 4 Q8NFO6	Q8NFO6 homo sapien
24	321	3.9	692 4 Q8ED51	Q8ED51 homo sapien
25	319	3.9	884 7 Q2P675	Q2P675 mus musculus
26	318	3.9	4123 4 Q7S551	Q7S551 homo sapien
27	317	3.9	1222 11 Q8K173	Q8K173 mus musculus
28	314	3.8	863 11 Q8C6J9	Q8C6J9 mus musculus
29	314	3.8	1464 11 Q8BK12	Q8BK12 mus musculus
30	313	3.8	1464 4 Q8N473	Q8N473 homo sapien
31	311	3.8	1637 6 Q8XSV8	Q8XSV8 bos taurus
32	310	3.8	1118 5 Q8YR13	Q8YR13 drosophila
33	308.5	3.8	1461 4 Q76045	Q76045 mus musculus
34	308	3.8	1453 11 Q63079	Q63079 rattus norv
35	307.5	3.7	977 7 Q8HNS9	Q8HNS9 mus musculus
36	307.5	3.7	1222 11 Q8K173	Q8K173 mus musculus
37	307	3.7	1314 11 Q8CGA5	Q8CGA5 mus musculus
38	305	3.7	397 11 Q8BUT6	Q8BUT6 mus musculus
39	304.5	3.7	1464 11 Q8BLM4	Q8BLM4 mus musculus
40	303	3.7	774 12 Q41971	Q41971 murid herpe

98	270	3.3	1442	11	Q62031	Q62031 mus musculus
99	270	3.3	1442	11	Q62032	Q62032 mus musculus
100	270	3.3	1459	11	Q62032	Q62032 mus musculus
101	269.5	3.3	608	12	Q95K9	Q95K9 herpessvirus
102	269.5	3.3	1453	11	Q63079	Q63079 rattus norv
103	269.5	3.3	1461	4	Q76045	Q76045 homo sapien
104	269.5	3.3	1464	4	Q8N473	Q8N473 homo sapien
105	268	3.2	998	11	Q8CTM4	Q8CTM4 mus musculus
106	267.5	3.2	1487	4	Q14047	Q14047 homo sapien
107	267	3.3	1458	13	Q91089	Q91089 oncorhynch
108	264.5	3.2	1160	4	Q14046	Q14046 homo sapien
109	264.5	3.2	1352	13	Q90Y10	Q90Y10 brachydanto
110	264.5	3.2	1838	4	Q15094	Q15094 homo sapien
111	263	3.2	894	5	Q8WMS4	Q8WMS4 mytilus gal
112	263	3.2	1419	11	Q63123	Q63123 rattus norv
113	263	3.2	1838	11	Q68207	Q68207 mus musculus
114	262	3.2	1445	13	Q93281	Q93281 rana catesb
115	262	3.2	1447	13	Q91B91	Q91B91 xenopus lae
116	262	3.2	1737	11	Q9J104	Q9J104 rattus norv
117	262	3.2	1835	13	Q91A04	Q91A04 gallus gall
118	261.5	3.2	1449	13	Q910C0	Q910C0 oncorhynch
119	261	3.2	1827	13	Q8U0M5	Q8U0M5 oryzias lat
120	260	3.1	620	16	Q8XZ99	Q8XZ99 rickettsia s
121	259.5	3.2	1491	13	Q91718	Q91718 xenopus lae
122	259	3.2	998	11	Q8CTM4	Q8CTM4 mus musculus
123	259	3.2	1838	11	Q88207	Q88207 mus musculus
124	258.5	3.1	1621	4	Q9H4F9	Q9H4F9 homo sapien
125	258	3.1	1349	13	Q8AW17	Q8AW17 brachydanto
126	257.5	3.1	1684	6	Q8HVC1	Q8HVC1 canis famli
127	257.5	3.1	1739	11	Q9J112	Q9J112 mus musculus
128	257	3.1	1258	13	Q8AM11	Q8AM11 brachydanto
129	256.5	3.1	926	5	Q9W3G1	Q9W3G1 drosophila
130	256.5	3.1	1491	13	Q91718	Q91718 xenopus lae
131	256.5	3.1	1838	4	Q15094	Q15094 homo sapien
132	256.5	3.1	1840	11	Q9J103	Q9J103 rattus norv
133	255.5	3.1	888	13	Q90796	Q90796 gallus gall
134	255.5	3.1	1739	11	Q9J112	Q9J112 mus musculus
135	255	3.1	748	11	Q99KMO	Q99KMO mus musculus
136	254	3.1	1347	4	Q9EQB3	Q9EQB3 homo sapien
137	253	3.1	566	6	Q95JDI	Q95JDI sus scrofa
138	253	3.1	1840	11	Q60467	Q60467 ericetulus
139	253	3.1	1840	11	Q60467	Q60467 ericetulus
140	252.5	3.1	2703	5	Q9VEG7	Q9VEG7 drosophila
141	252.5	3.1	2715	5	Q61603	Q61603 drosophila
142	252.5	3.1	2716	5	Q8N84	Q8N84 drosophila
143	251.5	3.1	3084	12	Q8U211	Q8U211 pseudorabie
144	251	3.0	771	4	Q9UJC7	Q9UJC7 homo sapien
145	251	3.1	890	5	Q77087	Q77087 alvirella p
146	250.5	3.1	1258	13	Q8AW11	Q8AW11 brachydanto
147	250.5	3.1	563	5	Q9B1T5	Q9B1T5 nephila mad
148	250.5	3.1	685	6	Q9RTS5	Q9RTS5 bos taurus
149	250.5	3.0	817	5	Q964V6	Q964V6 dictyosteli
150	250.5	3.0	1366	4	Q15177	Q15177 homo sapien

ALIGNMENTS

QY	181	TGTGAATGTCCTCGAGAGGCTTTTCAGGACAGAGGACCACTGCTGACCTGCTG	240
DB	6	CysAspMetCysSerClnGlnGlnuphneGlnAlaGlnArgSerGlnLeuValAlaLeuLeu	25
QY	241	GTCACAGGCTCCTCGAGGCTTCAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTG	300
DB	26	IlleSerGlnSerLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	45
QY	301	CTCTCCTGAGGACTACAGAGGCTTCACACTCTGAGGACAGGCTCTCTCCACTTGGCC	360
DB	46	LeuSerArgGlnUpSerPtyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	65
QY	361	AGGCGCTCTGAGACACCGCTGAGATAGGAGACTGGGCTCTGAGACTCATCGCG	420

RESULT 1  
Q6K320 PRELIMINARY; PRT; 1020 AA.  
ID Q6K320;  
AC Q6K320;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DI 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE CARD5.  
GN CARD5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BAH/c;  
RA Imanaga Y., Davey M.P., Martin T.M., Plonck S.R., DePriest M.L.,  
RT Baugh M.M., Sung C., Rosenbaum J.T.;  
RT "Cloning, sequencing and expression analysis of the murine Nod2/Card5  
gene";  
RL Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF520774; AA: M76073.1; --  
DR KGD; MG1:429387; Card5.  
DR InterPro; IPR001315; Card5.  
DR InterPro; IPR007091; LRR\_RN1nh.  
DR InterPro; IPR007111; NACHT\_NTPase.  
DR PROSITE; PS02099; CARD; 2\_DNase.  
DR PROSITE; PSS0503; LRR\_R1; 1.  
DR PROSITE; PSS0837; NACHT; 1.  
DR PROSITE; PSS137; TATD; 1; 1.  
SQ SEQUENCE 1020 AA; 113561 MW; 23504905ECF70EBB CRC64;

Alignment Scores:  
Pred. No.: 5,3e-289  
Score: 4196.00  
Percent Similarity: 87.398  
Best Local Similarity: 79.218  
Query Match: 50.748  
DB: 11  
Gaps: 0



Db 66 ArgArgLeuLeuAspThrValIlePheAsnLysGlyValIleProLysGlyLeuLeuLeu 85  
 QY 421 GCTGCCCAAGAAAGCCGAGCCCAACAGCCAGTCCCAAGTGCATGGCTGCTGGAGCCC 480  
 Db 86 AlaValGlnGlnAlaGlnAlaAsnSerHisThrPheLeuLeuTyrGlySerThrPhe 105  
 QY 481 CACTCGCTCCAGCCGAGCCGAGCTGAGAGTCAACCGGCAAGCCATTGTCAGAGGCTC 540  
 Db 106 HisSerLeuHisProThrThrGlyAspLeuSerHisSerProAlaIleValArgArgLeu 125  
 QY 541 CACAGCCATGTGAGAAACATGTGAACTGGAGTGGAGGGAGTTTCGACAGCCATAT 600  
 Db 126 TyrAsnHisValGlnAlaIleLeuGlnLeuAlaArgGlnGlyGlyPheLeuSerGlnTyr 145  
 QY 601 GAATGTGATGAATCAGGTTGGCCATCTTCACACCGTCCCAAGGGGCAAGAGGCTGTT 660  
 Db 146 GlnCysGlnGlnIleArgLeuProIlePheThrSerSerGlnArgAlaArgArgLeuLeu 165  
 QY 661 GATCTTGCCACGGTGAAGGCAATGATGGCTGCTTCCTTCTACAAATGTTCAAGAA 720  
 Db 166 AspLeuAlaIleValLysAlaAsnGlyLeuAlaAlaPheLeuLeuGlnHisValArgGln 185  
 QY 721 TTACCAATGCCATTGGCCCTGGCTTGGAAAGCTGCAATGCAGAGATATGGCCAA 780  
 Db 186 LeuProAlaProLeuProLeuProTyrGlnAlaAlaGlnCysGlnLysPheLeuSerLys 205  
 QY 781 CTGAGGACCAAGGTGTCTGCTCAGTCTCGTTCTCAGTACCTATGATGAGACAGAGCG 840  
 Db 206 LeuArgThrPheValLeuThrFelInsArgPheLeuSerThrTyrAspArgLysGlnAsn 225  
 QY 841 CTCTGCTGAGAGCATATACAGAGAGATGTCTGAGAGTCTGGGCAATGTGGGCAATG 900  
 Db 226 LeuCysLeuGlnAspIleTyrThrGlnAsnIleLeuLeuLeuGlnThrGlnValGlyThr 245  
 QY 901 GCTGGAACCCCGCAGAAAGAGCCAGCCAGCTGGGCTGGAGAGCTCTTCAGCAACCCCT 960  
 Db 246 AlaGlyAlaLeuGlnLysSerProAlaIleLeuGlnLysLeuGlnAspLeuHisThrHis 265  
 QY 961 GGCCAGCTCATGACATGGGAGACATGTGCTGGTGGTGGGTGAAGCGGAGCATGGGCAAG 1020  
 Db 266 GlnHisLeuAsnArgAspAlaAspThrIleLeuValValGlnGlnAlaGlySerGlnLys 285  
 QY 1021 AGCAGCTCTGCAAGCGGCTGCACTTGCTGTGGGCTGCAAGGCAAGCTTCAGAGAAATT 1080  
 Db 286 SerThrLeuLeuGlnArgLeuHisLeuLeuThrPheAlaThrGlyArgSerPheGlnGlnPhe 305  
 QY 1081 CTCTTGTCTTCCATTCAAGCTGCCGGAGCTGCAATGATGGCCAAACCACTCTCTGTG 1140  
 Db 306 LeuPheIlePheProPheSerCysArgGlnLeuGlnCysValAlaLysProLeuSerLeu 325  
 QY 1141 CGGACTCTACTCTTGAGACATGGCTGTGGGCTGATGTTGCTCAAGAAACATCTTCAG 1200  
 Db 326 ArgThrLeuLeuPheGlnHisCysCysTrpCysPheAlaAlaGlnAspArgAlaPheGln 345  
 QY 1201 TTACTCTGACCAACCTGACCGGTGCTCTTAACTTTGATGGCTTTCAGACAGTTCAAG 1260  
 Db 346 PheLeuLeuAspHisProAspArgValLeuLeuThrPheAspGlnLysLeuAspGlnPheLys 365

QY 1261 TTCAAGTTCACGATGCTGAAGGCACTGCTCCCGGACCGAACCCCACTCTGTCCAGACC 1320  
 Db 366 PheArgPheThrAspArgGlnArgHisCysSerProIleAspProThrSerValGlnThr 385  
 QY 1321 CTGCTCTCAACCTCTGCAAGGCGCAACCTGCTGAAGATCCCGGCAAGGTGGTACACAG 1380  
 Db 386 LeuLeuPheAsnLeuLeuGlnGlnLysAsnLeuLeuLysAsnAlaCysLysValLeuThrSer 405  
 QY 1381 CGTCCGCGCGCTGTGCGCGCTTCTCAAGAGATCATCCGACCAAGTTCAACTCAAG 1440  
 Db 406 ArgProAspAlaValSerAlaLeuLeuArgLysPheValArgThrGlnLeuGlnLeuLys 425  
 QY 1441 GGCCTCTGAAACAGGCGCATGAGCTGTCATCAGAGAGCCCATCATAGCCCGGAGTG 1500  
 Db 426 GlnPheSerGlnGlnLysIleGlnLeuLysArgLysHisSerGlnProGlyVal 445  
 QY 1501 GCGGACCGGCTCATCCGCGCTGCTCAAGAGACTCAGCCGCTGCAAGGTTGTGCCACTG 1560  
 Db 446 AlaAspArgLeuIleGlnLeuIleGlnAlaThrSerAlaLeuHisGlnLysLeuThrHis 465  
 QY 1561 CCGTCTCTCATGATGCTGTCAAATGCCACCAAGACTGTTGCTGAGAGAGGGGAG 1620  
 Db 466 ProValPheSerTrpMetValSerArgCysHisArgGlnLeuLeuGlnAsnArgGly 485  
 QY 1621 TCCCAAAAGACACATACAGATATGTAACGTGATCTGAGACATTTCTGCTGCAATGCC 1680  
 Db 486 PheProThrThrSerThrAspMetLysLeuLeuIleLeuGlnHisPheLeuLeuHisAla 505  
 QY 1681 ACCCCGCCAATCAAGCTTCCCAAGGTCTGGAGCCCAAGTCTTTCGGGGCCGCTCCCC 1740  
 Db 506 SerProProAspSerSerProLeuGlnLysGlyProGlnLysLeuGlnInsArgLeuSer 525  
 QY 1741 ACCCTCTGCAACCTGGGCGAGACTGGCTGTGGGCGCTGGGCAATGTGCTACGTGTC 1800  
 Db 526 ThrLeuLeuHisLeuGlnHisLeuAlaLeuArgGlyLeuAlaIleSerCysTyrValPhe 545  
 QY 1801 TCAAGCCAGAGCTCCAGAGCAGCAGAGTCAAGCCCTGATGACATTTCTCTGGCTTCTG 1860  
 Db 546 SerAlaGlnGlnLeuGlnAlaAlaGlnValAspSerAspAspIleSerLeuGlnPheLeu 565  
 QY 1861 GTGCGTGCMAAAGTGTGCTGCCAGGAGTACGGCCCGCTGGAATTCCTTCACATCACT 1920  
 Db 566 ValArgAlaGlnSerSerValProGlySerLysAlaProLeuGlnPheLeuHisIleThr 585  
 QY 1921 TTCAGTCTCTTTGCGCGCTTCTACTGAGCACTCAGTGTGATGTGCACACAGCTTGG 1980  
 Db 586 PheGlnCysPhePheAlaAlaPheTyrLeuAlaValSerAlaAspThrSerValAlaSer 605  
 QY 1981 CTCAGACACTCTTCAATGTGGCAGCCGAGGCACTCACCAATGGCAGGCTCTGCC 2040  
 Db 606 LeuLysHisLeuPheSerCysGlnYArgGlnGlySerSerLeuLeuGlnYArgLeuLeuPro 625  
 QY 2041 ACCATGTGATCCAGGCTCGAGAGGAAAGAGCAGCGTGGCAACTTCTGCGCAAG 2100  
 Db 626 AsnLeuCysIleGlnLysSerArgValLysLysGlnSerGlnAlaAlaLeuLeuGlnLys 645

